Interference Alignment: From Degrees-of-Freedom to Constant-Gap Capacity Approximations

Urs Niesen and Mohammad A. Maddah-Ali

arXiv:1112.4879v1 [cs.IT] 20 Dec 2011

Abstract

Interference alignment is a key technique for communication scenarios with multiple interfering links. In several such scenarios, interference alignment was used to characterize the degrees-of-freedom of the channel. However, these degrees-of-freedom capacity approximations are often too weak to make accurate predictions about the behavior of channel capacity at finite signal-to-noise ratios (SNRs). The aim of this paper is to significantly strengthen these results by showing that interference alignment can be used to characterize capacity to within a constant gap. We focus on real time-invariant frequency-flat X-channels. The only known solutions achieving the degrees-of-freedom of this channel are either based on real interference alignment or on layer-selection schemes. Neither of these solutions seems sufficient for a constant-gap capacity approximation.

In this paper, we propose a new communication scheme and show that it achieves the capacity of the Gaussian X-channel to within a constant gap. To aid in this process, we develop a novel deterministic channel model. This deterministic model depends on the $\frac{1}{2}\log(\mathsf{SNR})$ most-significant bits of the channel coefficients rather than only the single most-significant bit used in conventional deterministic models. The proposed deterministic model admits a wider range of achievable schemes that translate to a solution for the Gaussian channel. For this deterministic model, we find an approximately optimal communication scheme. We then translate the solution for the deterministic channel to the original Gaussian X-channel and show that it achieves capacity to within a constant gap.

I. Introduction

Interference alignment has been shown to be a key technique to achieve optimal degrees-of-freedom (capacity pre-log factor) in several common wireless network configurations such as X-Channels [1], interference channels [2], [3], interfering multiple-access and broadcast channels [4], multi-user systems with delayed feedback [5]–[7], and distributed computation [8], among others. The main idea of interference alignment is to force all interfering signals at the receivers to be aligned, thereby maximizing the number of interference-free signaling dimensions.

A. Background

Alignment approaches can be divided into two broad categories.

1) *Vector-space alignment [1], [2]*: In this approach, conventional communication dimensions, such as time, frequency, and transmit/receive antennas, are used to align interference. At the transmitters, precoding matrices are designed over multiple time/frequency slots or transmit antennas such that interference at the receivers is aligned in a small subspace. At the receivers, zero-forcing is used to remove this subspace containing the interference. If the channel coefficients have enough variation across the utilized time/frequency slots or antennas, then such precoding matrices can be found.
2) *Signal-scale alignment [3], [9]*: If the transmitters and receivers have only a single antenna and the channel coefficients are time invariant and frequency flat, the vector-space alignment method fails. Instead, one can make use of another resource, namely is the signal scale. Using lattice codes, the transmitted and received signals are split into several superimposed layers. The transmitted signals are chosen such that all interfering signals are observed within the same layers at the receivers. Thus, alignment is now achieved in signal scale.

The authors are with Bell Labs, Alcatel-Lucent. Emails: urs.niesen@alcatel-lucent.com, mohammadali.maddah-ali@alcatel-lucent.com

Signal-scale interference alignment can be further subdivided into two different, and seemingly completely unrelated, approaches: alignment schemes motivated by *signal-strength deterministic models* [9], [10] and *real interference alignment* [3].

For the deterministic approach, the channel is first approximated by a deterministic noise-free channel. In this deterministic model, all channel inputs and outputs are binary vectors, representing the binary expansion of the real valued signals in the Gaussian case. The actions of the channel are modeled by shifting these vectors up or down, depending on the most-significant bit of the channel gains, and by bitwise addition over the binary field of interfering vectors. The signal layers are represented by the different bits in the binary expansion of the signals. In the second step, the signaling schemes and the outer bounds developed for this simpler deterministic model are used to guide the design of efficient signaling schemes for the original Gaussian problem.

This deterministic approach has proved instrumental in approximating the capacity of several challenging multi-user communication networks. Using this approach, constant-gap capacity approximations were derived for example for single-multicast relay networks [11], two-user interference channels with feedback [12] or with transmit/receive cooperation [13], [14], and lossy distributed source coding [15]. In all these communication scenarios, interference alignment is not required. For communication scenarios in which interference alignment is required, the deterministic approach has been less helpful. In fact, it has only been successfully used to obtain constant-gap capacity approximations for the fairly restrictive many-to-one interference channel, in which only one of the receivers experiences interference while all others are interference free [9]. Even for the X-channel, the simplest Gaussian network in which interference alignment is required, only weaker (generalized) degrees-of-freedom capacity approximations were derived using the deterministic approach [10]. The resulting communication scheme for the Gaussian X-channel in [10] is quite complicated and can not be used to derive a constant-gap capacity approximation.

For the real interference alignment approach, each transmitter modulates its signal using an integer lattice transmit constellation. The transmit constellations are then scaled by a constant to ensure that at each receiver all interfering lattice constellations coincide, while desired lattice constellations are disjoint. Each receiver recovers the desired signals using a minimum-distance decoder. A number-theoretic result concerning the approximability of real numbers by rationals, called Groshev's theorem, is used to analyze the minimum constellation distance at the receivers. For almost all channel gains, this scheme is shown to achieve the full degrees-of-freedom of the Gaussian X-channel and the Gaussian interference channel [3]. While this scheme is asymptotically optimal for almost all channel gains, there are infinitely many channel gains for which the scheme fails, for example when the channel gains are rational. Moreover, this approach can again not be used to derive stronger constant-gap capacity approximations.

At first glance, real interference alignment appears to rely on the irrationally of the channel coefficients, preventing the desired integer input signals from mixing with the undesired integer interference signals. This raises the concern that the scheme might be severely affected by the presence of measurement errors or quantization of the channel coefficients. In addition, arbitrarily close to any irrational channel realization is a rational channel realization. How are we then to engineer a communication device based on this scheme? Quoting from Slepian's 1974 Shannon Lecture [16]: "Most of us would treat with great suspicion a model that predicts stable flight for an airplane if some parameter is irrational but predicts disaster if that parameter is a nearby rational number. Few of us would board a plane designed from such a model."

Some of these concerns follow from the fact that real interference alignment is somehow isolated from other known signaling schemes and only poorly understood. Unlike the vector-space and the deterministic approaches, no vector space interpretation is known for real interference alignment, making it harder to obtain intuition. On the other hand, it is known that the degrees-of-freedom of the interference channel itself is an everywhere discontinuous function of the channel realizations [17]. It should therefore not be surprising that the rates achieved by real interference alignment share this characteristic. Rather, it appears that the degrees-of-freedom capacity approximation itself is too weak to make accurate predictions about the behavior of channel capacity at finite SNRs, and that the everywhere discontinuity of the degrees-of-

freedom in the channel coefficients are mainly caused by taking a limit as SNR approaches infinity. Thus, a stronger capacity approximation is needed.

B. Summary of Results

In this paper, we derive such a stronger capacity approximation for the Gaussian X-channel by characterizing its capacity to within a constant gap independent of the SNR and the channel gains, up to an (easily computable) "outage" set of channel realizations of arbitrarily small measure. In order to obtain this approximation result, we develop a novel deterministic channel model. In this deterministic model, each channel gain is modeled by a lower-triangular binary Toeplitz matrix. The entries in this matrix consist of the first $\frac{1}{2}\log(\mathsf{SNR})$ bits in the binary expansion of the channel gain in the corresponding Gaussian model. This contrasts with the traditional deterministic model, which is based only on the *single* most significant nonzero bit in the binary expansion of the Gaussian channel gain. The proposed deterministic model is rich enough to explain the real interference alignment approach. Thus, it unites the so far disparate deterministic and real interference alignment approaches mentioned above. Moreover, as our proposed deterministic model is based on a vector space, it enables an intuitive vector space interpretation of real interference alignment. This deterministic model allows us to design a fairly simple signaling scheme that achieves capacity of the deterministic channel up to a constant gap, as long as the binary channel matrices satisfy certain rank conditions.

The solution for the proposed deterministic model is simple enough to be mimicked for the original Gaussian channel. We show that this scheme achieves the capacity of the Gaussian X-channel to within a constant gap. To prove achievability for the Gaussian case, we extend Groshev's theorem to handle finite SNRs as well as channel gains of different magnitudes. To derive the constant-gap approximation of capacity, we introduced a new outer bound on the sum capacity of the Gaussian X-channel. In addition, we show that, similarly to the Gaussian point-to-point channel, capacity is not sensitive to quantization and measurement errors in the channel gains smaller than $\mathsf{SNR}^{-1/2}$.

One implication of these result is that the complicated solution achieving the degrees-of-freedom of the Gaussian X-channel in [10] is a result of oversimplification in the signal-strength deterministic model rather than the properties of the original Gaussian channel itself. Moreover, the results in this paper imply that the discontinuity of the degrees-of-freedom of the Gaussian X-channel with respect to the channel coefficients is due to the large SNR limit and is not present at finite SNRs.

C. Organization

The remainder of this paper is organized as follows. Section II introduces the new deterministic channel model. Section III formalizes the Gaussian network model and the problem statement. Section IV presents the main results of the paper—Sections V and VI contain the corresponding proofs. Section VII contains the mathematical foundations for the analysis of the decoding algorithms. Section VIII concludes the paper.

II. Deterministic Channel Models

Developing capacity achieving communication schemes for multi-user communication networks is often challenging. Indeed, even for the relatively simple two-user interference channel, finding capacity is a long-standing open problem. For the Gaussian network, the difficulty is due to the interaction between the various components of these networks, such as broadcast, multiple access, and additive noise. For example, the two-user interference channel mentioned before has two broadcast links, two multiple-access links, and two additive noise components.

Surprisingly, some results in the field of network information theory suggest that if the noise components are eliminated, so that the channel output at the receivers becomes a deterministic function of the channel inputs at the transmitters, then the problem of characterizing capacity can be substantially simplified [11],

[18]. Such networks are called deterministic networks. This observation motivates the investigation of noisy networks by approximating them with deterministic networks [11], [19], [20].

This approximation has two potential advantages. First, the capacity of the deterministic network may directly approximate the capacity of the original Gaussian network. Second and more important, the deterministic model may reveal the essential ingredients of an efficient signaling scheme for the noisy network. In other words, the capacity achieving signaling scheme for the deterministic network may be used as a road map to design signaling schemes for the Gaussian network. If the deterministic approximation is well chosen, then the resulting signaling scheme for the Gaussian network is close to capacity achieving.

The first critical step in this approach is thus to find an appropriate deterministic channel approximating the Gaussian one. This deterministic channel model should satisfy two criteria.

1) *Simplicity*: The deterministic model should be simple to deal with. Obviously, we do not want to transform a difficult Gaussian problem into an equally difficult deterministic one.
2) *Richness*: The deterministic model should keep the main features of the original problem. If the problem is oversimplified, then the solution for the deterministic model may not lead to a good solution for the Gaussian problem.

These two requirements are conflicting. Indeed, oversimplification of the Gaussian model can sacrifice the richness of the deterministic model. Conversely, keeping too many of the features of the Gaussian model can result in a deterministic model that is rich but too difficult to analyze. Striking the right balance between these two requirements is the key to developing a useful deterministic network approximation.

One of the approaches that achieves this goal is the signal-strength deterministic model proposed by Avestimehr et al. [11]. We review this deterministic model in Section II-A. We introduce our new lower-triangular deterministic model in Section II-B. Section II-C compares the two deterministic models, explaining the shortcomings of the former and the need for the latter.

A. Signal-Strength Deterministic Model [11]

We start with the real point-to-point Gaussian channel

$$y[t] \triangleq 2^n h x[t] + z[t], \tag{1}$$

with additive white Gaussian noise $z[t] \sim \mathcal{N}(0,1)$ and unit average power constraint

$$\frac{1}{T}\sum_{t=1}^{T} x^2[t] \leq 1.$$

Here, n is a positive integer, and $h \in [1,2)$. Observe that all channel gains (and hence SNRs) greater than or equal to one can be expressed in the form $2^n h$ for n and h satisfying these conditions. If the magnitude of the channel gains is less than one, then capacity is less than one bit per channel use and hence not relevant for capacity approximation up to a constant gap. Moreover, since capacity is only a function of the magnitude of the channel gains, negative channel gains are not relevant here. Hence, (1) is essentially the general case.

To develop the deterministic model and for simplicity, we assume that $x[t]$ and $z[t]$ are positive and upper bounded by one. We can then write x and z in terms of their binary expansion as

$$x = \sum_{i=1}^{\infty} [x]_i 2^{-i} = 0.[x]_1[x]_2[x]_3\dots, \tag{2a}$$

$$z = \sum_{i=1}^{\infty} [z]_i 2^{-i} = 0.[z]_1[z]_2[z]_3\dots. \tag{2b}$$

The Gaussian point-to-point channel (1) can then be approximated as

$$\begin{aligned} y &= \sum_{j=-\infty}^{\infty} [y]_j 2^{-j} \\ &\approx 2^n x + z \\ &= \sum_{j=1}^{n} [x]_j 2^{n-j} + \sum_{j=1}^{\infty} ([x]_{j+n} + [z]_j) 2^{-j} \\ &\approx \sum_{j=1}^{n} [x]_j 2^{n-j}, \end{aligned}$$

or, more succinctly,

$$[y]_{j-n} \approx [x]_j, \text{ for } 1 \leq j \leq n,$$

see Fig. 1(a).

The approximation in this derivation is to ignore the impact of $h \in [1, 2)$, the noise, as well as all bits $[x]_{n+1}, [x]_{n+2}, \ldots$ with exponent less than zero. These bits with exponent less than zero are approximated as being completely corrupted by noise, whereas the bits with higher exponent are approximated as being received noise free. Therefore, we can approximate the Gaussian channel with a deterministic channel consisting of n parallel error-free links from the transmitter to the receiver, each carrying one bit per channel use.

Having reviewed the signal-strength model for the point-to-point case, we now turn to the Gaussian multiple-access channel. In addition to the noise components in the point-to-point model, this channel model also incorporates signal interaction. Consider the Gaussian multiple-access channel

$$y[t] \triangleq 2^n h_1 x_1[t] + 2^n h_2 x_2[t] + z[t], \tag{3}$$

where $z[t] \sim \mathcal{N}(0, 1)$ is additive white Gaussian noise. As before, we impose a unit average transmit power constraint on $x_1[t]$ and $x_2[t]$. Moreover, n is a nonnegative integer, and $h_1, h_2 \in [1, 2)$. The signal-strength deterministic model corresponding to the Gaussian channel (3) is

$$[y]_{j-n} \approx [x_1]_j \oplus [x_2]_j, \text{ for } 1 \leq j \leq n, \tag{4}$$

where $\oplus$ denotes addition over $\mathbb{Z}_2$, i.e., modulo two.

We note that in this model (i) the contribution of h is entirely ignored; (ii) real addition is replaced with bit-wise modulo two addition; (iii) noise is eliminated. As mentioned earlier, this simple model has been used to characterize the capacity region of several challenging problems in network information theory to within a constant gap. However, it falls short for some other settings. For example, for certain relay networks with specific channel parameters, this model incorrectly predicts capacity zero. Similarly, for interference channels with more than two users and for X-channels, this model fails to predict the correct behavior for the Gaussian case.

B. Lower-Triangular Deterministic Model

The signal-strength deterministic model, recalled in the last section, ignores the contribution of $h \in [1, 2)$ in the Gaussian point-to-point channel (1). Indeed, h is approximated by 1. In this section, we introduce a new deterministic channel model, termed *lower-triangular deterministic model*, in which the effect of h is preserved. As we will see later, the new deterministic model admits a wider range of solutions—a fact that will be critical for the approximation of Gaussian networks with multiple interfering signals.

Consider again the Gaussian point-to-point channel (1). Write the channel parameter $h \in [1, 2)$ in terms of its binary expansion

$$h = \sum_{j=0}^{\infty} [h]_j 2^{-j} = [h]_0.[h]_1[h]_2[h]_3 \ldots .$$

(a) Signal-strength deterministic model (b) Lower-triangular deterministic model

Fig. 1. Comparison of the signal-strength deterministic model [11], and the lower-triangular deterministic model proposed in this paper. In the figure, solid lines depict noiseless binary links of capacity one bit per second. Dashed lines depict noiseless links of either capacity one or zero bits per second (depending on whether the corresponding entry in the channel matrix $\bar{\boldsymbol{H}}$ is one or zero). Links with the same color/shade have the same capacity.

Observe that $[h]_0 = 1$, due to the assumption that $h \in [1, 2)$. Then, from (1), and (2), we have

$$\begin{aligned} y &= \sum_{j=-\infty}^{\infty} [y]_j 2^{-j} \\ &= 2^n \sum_{j=0}^{\infty} [h]_j 2^{-j} \sum_{i=1}^{\infty} [x]_i 2^{-i} + \sum_{j=1}^{\infty} [z]_j 2^{-j} \\ &= \sum_{j=1}^{n} \sum_{i=1}^{j} [h]_{j-i} [x]_i 2^{n-j} + \sum_{j=1}^{\infty} \Big(\sum_{i=1}^{j+n} [h]_{j+n-i} [x]_i + [z]_j \Big) 2^{-j} \\ &\approx \sum_{j=1}^{n} \Big(\sum_{i=1}^{j} [h]_{j-i} [x]_i \Big) 2^{n-j}, \end{aligned}$$

so that

$$[y]_{j-n} \approx \sum_{i=1}^{j} [h]_{j-i} [x]_i, \text{ for } 1 \leq j \leq n.$$

The approximation here is to ignore the noise as well as all bits in the convolution of $[h]_0.[h]_1[h]_2\dots$ and $0.[x]_1[x]_2\dots$ with exponent less than zero. These bits with exponent less than zero are approximated as being completely corrupted by noise, whereas the bits with higher exponent are approximated as being received noise free.

This suggests to approximate the Gaussian point-to-point channel (1) by a deterministic channel between the binary input vector

$$\bar{\boldsymbol{x}} \triangleq \begin{pmatrix} \bar{x}_1 & \bar{x}_2 & \dots & \bar{x}_n \end{pmatrix}^{\mathsf{T}}$$

and the binary output vector

$$\bar{\boldsymbol{y}} \triangleq \begin{pmatrix} \bar{y}_1 & \bar{y}_2 & \dots & \bar{y}_n \end{pmatrix}^{\mathsf{T}}$$

connected through the channel operation

$$\bar{\boldsymbol{y}} \triangleq \bar{\boldsymbol{H}} \bar{\boldsymbol{x}}, \tag{5}$$

with

$$\bar{\boldsymbol{H}} \triangleq \begin{pmatrix} [h]_0 & 0 & \cdots & 0 & 0 \\ [h]_1 & [h]_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ [h]_{n-2} & [h]_{n-3} & \cdots & [h]_0 & 0 \\ [h]_{n-1} & [h]_{n-2} & \cdots & [h]_1 & [h]_0 \end{pmatrix},$$

(a) Signal-strength deterministic model   (b) Lower-triangular deterministic model

Fig. 2. Permissible signaling schemes for both deterministic models.

(a) Signal-strength deterministic model   (b) Lower-triangular deterministic model

Fig. 3. Illustration of a signaling scheme that succeeds for the lower-triangular model (assuming the subspace condition (7) holds), but fails for the signal-strength model.

as depicted in Fig. 1(b). Here, we have normalized the received vector $\bar{\boldsymbol{y}}$ to contain the bits from 1 to n. This is a deterministic channel with finite input and output alphabets. Note that all operations in (5) are over $\mathbb{Z}_2$, i.e., modulo two. Similarly, the Gaussian multiple-access channel (3) can be approximated by the deterministic channel model

$$\bar{\boldsymbol{y}} \triangleq \bar{\boldsymbol{H}}_1 \bar{\boldsymbol{x}}_1 \oplus \bar{\boldsymbol{H}}_2 \bar{\boldsymbol{x}}_2. \tag{6}$$

C. Comparison of Deterministic Models

We now compare the signal-strength deterministic model reviewed in Section II-A and the lower-triangular deterministic model introduced in Section II-B. As an example, we consider the Gaussian multiple-access channel (3) with signal strength $n = 4$. The corresponding deterministic models are given by (4) and (6). Assume that transmitter one wants to send three bits a_1, a_2, and a_3 to the receiver. At the same time, transmitter two wants to send one bit b_1.

In both deterministic models (4) and (6), transmitter one can use the first three layers to send a_1, a_2, and a_3, while transmitter two can use the last layer to send b_1, as shown in Fig. 2. For the signal-strength model, the decoding scheme is trivial. For the lower-triangular model, the receiver starts by decoding the highest layer containing only a_1. Having recovered a_1, the receiver cancels out its contribution in all lower layers. The decoding process continues in the same manner with a_2 at the second-highest layer, until all bits are decoded.

There are, however, some signaling schemes that are only decodable in the lower-triangular model, but not in the signal-strength model. An example of such a signaling scheme is depicted in Fig. 3. In this

scheme, transmitter one uses again the first three layers to send a_1, a_2, and a_3. Unlike before, transmitter two now also uses the first layer to send b_1. From Fig. 3(a), we can see that, in the signal-strength model, receiver one observes $a_1 \oplus b_1$ and cannot recover a_1 and b_1 from the received signal. However, this scheme can be utilized successfully in the lower-triangular model as long as the subspaces spanned by the message bits at the receivers are linearly independent. In this case, the subspace spanned by the first three columns of $\bar{\boldsymbol{H}}_1$ and the subspace spanned by the first columns of $\bar{\boldsymbol{H}}_2$ need to be linearly independent. This is the case if and only if

$$\det \begin{pmatrix} [h_1]_0 & 0 & 0 & [h_2]_0 \\ [h_1]_1 & [h_1]_0 & 0 & [h_2]_1 \\ [h_1]_2 & [h_1]_1 & [h_1]_0 & [h_2]_2 \\ [h_1]_3 & [h_1]_2 & [h_1]_1 & [h_2]_3 \end{pmatrix} \neq 0. \tag{7}$$

The event (7) depends not only on n, but also on the bits in the binary expansion of h_1 and h_2. Thus, this scheme is successful for all channel gains $(h_1, h_2) \in (1, 2]^2 \setminus B$, where B is the event that (7) does not hold. The set B can be understood as an outage event: If the channel gains are in B, the achievable scheme fails to deliver the desired target rate of 4 bits per channel use.

Noting that the scheme depicted in Fig. 2(b) always works while the scheme depicted in Fig. 3(b) only works under some conditions, one might question the relevance of the new class of solutions. The answer is that it is precisely this class of signaling solutions that are valid for the lower-triangular model but invalid for the signal-strength model that are efficient for the X-channel to be investigated in Section IV.

As pointed out earlier, the second step in using the deterministic approach is to transfer the solution for the deterministic model to a solution for the original Gaussian model. We now show how this can be done for the signaling scheme shown in Fig. 3(b). The proposed scheme for the Gaussian multiple-access channel is depicted in Fig. 4. In this scheme, the input constellation at transmitter one is the set $\{0, 1/8, \dots, 7/8\}$, and the input constellation at transmitter two is the set $\{0, 1/2\}$. Since the additive Gaussian receiver noise has unit variance, we expect the receiver to be able to recover the coded input signals roughly when

$$2^n |h_1 u_1 + h_1 u_2 - h_1 u_1' - h_2 u_2'| > 2 \tag{8}$$

for all $u_1, u_1' \in \{0, 1/8, \dots, 7/8\}$, $u_2, u_2' \in \{0, 1/2\}$ such that $(u_1, u_2) \neq (u_1', u_2')$. In words, we require the minimum constellation distance as seen at the receiver to be greater than two.

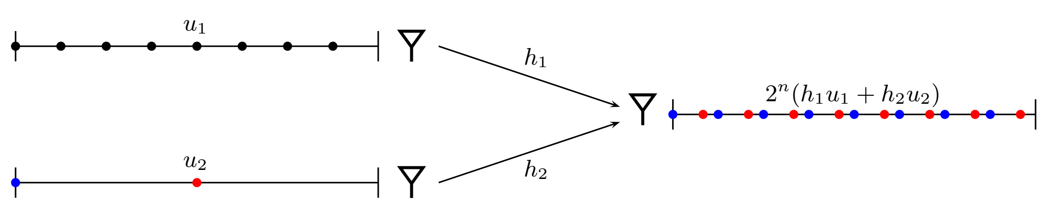

Fig. 4. Modulation scheme for the Gaussian model suggested by the signaling scheme for the lower-triangular deterministic model depicted in Fig. 3(b). At the decoder, blue dots correspond to input tuples $(u_1, 0)$ with $u_1 \in \{0, 1/8, \dots, 7/8\}$, and red dots correspond to input tuples $(u_1, 1/2)$ with $u_1 \in \{0, 1/8, \dots, 7/8\}$. Here, $n = h_1 = 1$ and $h_2 = 1/6$.

We note that condition (8) for the Gaussian channels corresponds to condition (7) for the deterministic model. As in the deterministic case, this scheme fails to work whenever the channel gains are in the set B not satisfying (8), and one can bound the Lebesgue measure of this outage event B. It is worth emphasizing that condition (8) has nothing to do with the rationality or irrationality of the channel coefficients.

Remark: In the special case in which each transmitter has the same message size, the modulation scheme shown in Fig. 4 is the same as the modulation scheme used in *real interference alignment* [3], [17]. The objective in [3] is to achieve only the degrees-of-freedom of the channel, and therefore the modulation scheme there is designed and calibrated for the high SNR regime. As a result, the modulation scheme in [3] is not sufficient to prove a constant-gap capacity approximation. Rather, as we will see

in Section IV, asymmetric message sizes and judicious layer selection guided by the proposed lower-triangular deterministic model together with a more careful and more general analysis of the receivers are required to move from the degrees-of-freedom to the constant-gap capacity approximation.

III. Network Model

In the remainder of the paper, we focus on the X-channel. This section formally introduces the channel model and capacity definitions. We start by introducing notational conventions in Section III-A. We introduce the symmetric X-channel in Section III-B, and the general asymmetric X-channel in Section III-C.

A. Notation

Throughout this paper, we use small and capital bold font to denote vectors and matrices, i.e., $\boldsymbol{x}$ and $\boldsymbol{H}$. For a real number $a \in \mathbb{R}$, we use $(a)^+$ to denote $\max\{a, 0\}$. For a set $B \subset \mathbb{R}^d$, $\mu(B) = \mu_d(B)$ denotes d-dimensional Lebesgue measure. For an event A, $\mathbb{1}_A$ denotes the indicator function of A. Finally, all logarithms are expressed to the base two. As a result, all capacities are expressed in bits per channel use.

B. X-Channel with Symmetric SNR*s*

The Gaussian X-channel with symmetric SNRs consists of two transmitters and two receivers. The *channel output* $y_m[t]$ at receiver $m \in \{1, 2\}$ and time $t \in \mathbb{N}$ is

$$y_m[t] \triangleq 2^n h_{m1} x_1[t] + 2^n h_{m2} x_2[t] + z_m[t], \tag{9}$$

where $x_k[t]$ is the *channel input* at transmitter $k \in \{1, 2\}$, where h_{mk} is the (constant) *channel gain* from transmitter k to receiver m, and where $z_m[t] \sim \mathcal{N}(0, 1)$ is additive white Gaussian *receiver noise*. The channel gains consist of two parts, 2^n and h_{mk}. We assume here that $n \in \mathbb{Z}_+$, and that $h_{mk} \in (1, 2]$ for each m, k. Thus we see that 2^{2n} captures the approximate SNR and h_{mk} captures the fine structure of the channel gains. Since each channel gain is multiplied by the same number 2^n, each of the four links has the same approximate SNR—the general case with asymmetric SNRs is discussed in the next section.

Each transmitter has one *message* to communicate to each receiver. So there are a total of four mutually independent messages w_{mk} with $m, k \in \{1, 2\}$. The message w_{mk} is uniformly distributed in $\{1, \dots, W_{mk}\}$. An *encoder* of *block length* T at transmitter k is a function

$$f_k : \{1, \dots, W_{1k}\} \times \{1, \dots, W_{2k}\} \to \mathbb{R}^T$$

mapping the messages w_{1k} and w_{2k} into the sequence of channel inputs

$$(x_k[t])_{t=1}^T \triangleq f_k(w_{1k}, w_{2k}).$$

The *rate* of the encoder f_k for message w_{mk} is

$$R_{mk} \triangleq \frac{\log(W_{mk})}{T}.$$

The encoder is said to satisfy a *unit average power constraint* if

$$\frac{1}{T} \sum_{t=1}^{T} x_k^2[t] \leq 1$$

for every w_{1k}, w_{2k}. A *decoder* of block length T at receiver m is a function

$$\phi_m : \mathbb{R}^T \to \{1, \dots, W_{m1}\} \times \{1, \dots, W_{m2}\}$$

mapping the sequence of channel outputs $(y_m[t])_{t=1}^T$ into the *message estimates*

$$(\hat{w}_{m1}, \hat{w}_{m2}) \triangleq \phi_m\big((y_m[t])_{t=1}^T\big)$$

of the messages (w_{m1}, w_{m2}). The *probability of error* of a pair of encoders (f_1, f_2) and decoders (ϕ_1, ϕ_2) is

$$\mathbb{P}\Big(\bigcup_{m,k}\{\hat{w}_{mk} \neq w_{mk}\}\Big).$$

Definition. A sum rate $R(n)$ is *achievable*, if for every $\varepsilon > 0$ and every $T \geq T_0(\varepsilon)$ there exists encoders (f_1, f_2) with block length T, unit average power constraint, and rates $\sum_{m,k} R_k \geq R(n)$, and there exists decoders (ϕ_1, ϕ_2) with block length T and probability of error less than ε. The *sum capacity* $C(n)$ of the Gaussian X-channel (9) is the supremum of all achievable sum rates $R(n)$.

In the following, we will be interested in a particular modulation scheme for the Gaussian channel, which we describe next. Fix a time slot t; to simplify notation, we will drop the dependence of variables on t whenever there is no risk of confusion. Assume each message w_{mk} is modulated into the signal u_{mk} of the form

$$u_{mk} \triangleq \sum_{i=3}^{\infty} [u_{mk}]_i 2^{-i} \tag{10}$$

with $[u_{mk}]_i \in \{0, 1\}$. Transmitter one forms the channel input

$$x_1 \triangleq h_{22}u_{11} + h_{12}u_{21}. \tag{11a}$$

Similarly, transmitter two forms the channel input

$$x_2 \triangleq h_{11}u_{22} + h_{21}u_{12}. \tag{11b}$$

Since $h_{mk}^2 \leq 4$ and $u_{mk}^2 \leq 1/16$, this satisfies the unit average power constraint at the transmitters.

The channel output at receiver one is then

$$\begin{aligned} y_1 &= 2^n h_{11}x_1 + 2^n h_{12}x_2 + z_1 \\ &= 2^n h_{11}h_{22}u_{11} + 2^n h_{12}h_{21}u_{12} + 2^n h_{11}h_{12}(u_{21} + u_{22}) + z_1, \end{aligned} \tag{12}$$

and similar at receiver two. Receiver one is interested in the signals u_{11} and u_{12}. The other two signals u_{21} and u_{22} are interference. We see from (12) that the interfering signals u_{21} and u_{22} are received with the same coefficient $h_{11}h_{12}$.

It will be convenient in the following to refer to the effective channel gains including the modulation scheme as g_{mk}, i.e.,

$$g_{10} \triangleq h_{11}h_{12}, \qquad g_{20} \triangleq h_{22}h_{21}, \tag{13a}$$

$$g_{11} \triangleq h_{11}h_{22}, \qquad g_{21} \triangleq h_{21}h_{12}, \tag{13b}$$

$$g_{12} \triangleq h_{12}h_{21}, \qquad g_{22} \triangleq h_{22}h_{11}. \tag{13c}$$

Here g_{mk} for $m, k \in \{1, 2\}$ corresponds the desired signal u_{mk}, and g_{m0} for $m \in \{1, 2\}$ corresponds to the interference terms. Since $h_{mk} \in (1, 2]$, we have $g_{mk} \in (1, 4]$.

As in the discussion in Section II-B, it is insightful to consider the lower-triangular deterministic equivalent of the modulated Gaussian X-channel (12). To simplify the discussion, we assume for the derivation and analysis of the deterministic channel model that the channel gains g_{mk} are in $(1, 2]$—the Gaussian setting will be analyzed for the general case. Let

$$\bar{\boldsymbol{G}}_{mk} \triangleq \begin{pmatrix} [g_{mk}]_0 & 0 & \cdots & 0 & 0 \\ [g_{mk}]_1 & [g_{mk}]_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ [g_{mk}]_{n-2} & [g_{mk}]_{n-3} & \cdots & [g_{mk}]_0 & 0 \\ [g_{mk}]_{n-1} & [g_{mk}]_{n-2} & \cdots & [g_{mk}]_1 & [g_{mk}]_0 \end{pmatrix} \tag{14}$$

be the deterministic channel matrix corresponding to the binary expansion of the channel gain g_{mk} with $m \in \{1,2\}$ and $k \in \{0,1,2\}$. Since $g_{mk} \in (1,2]$ by assumption, the diagonal values of $\bar{\boldsymbol{G}}_{mk}$ are equal to one. The lower-triangular deterministic equivalent of the modulated Gaussian X-channel is then given by

$$\bar{\boldsymbol{y}}_1 \triangleq \bar{\boldsymbol{G}}_{11}\bar{\boldsymbol{u}}_{11} \oplus \bar{\boldsymbol{G}}_{12}\bar{\boldsymbol{u}}_{12} \oplus \bar{\boldsymbol{G}}_{10}(\bar{\boldsymbol{u}}_{21} \oplus \bar{\boldsymbol{u}}_{22}), \tag{15a}$$

$$\bar{\boldsymbol{y}}_2 \triangleq \bar{\boldsymbol{G}}_{22}\bar{\boldsymbol{u}}_{22} \oplus \bar{\boldsymbol{G}}_{21}\bar{\boldsymbol{u}}_{21} \oplus \bar{\boldsymbol{G}}_{20}(\bar{\boldsymbol{u}}_{12} \oplus \bar{\boldsymbol{u}}_{11}), \tag{15b}$$

where the *channel input* $\bar{\boldsymbol{u}}_{mk}$ and the *channel output* $\bar{\boldsymbol{y}}_m$ are all binary vectors of length n, and where all operations are over $\mathbb{Z}_2$.[1]

As before, transmitter one has access to the *messages* w_{11}, w_{21}, and transmitter two has access to w_{12}, w_{22}. The four messages are mutually independent, and w_{mk} is uniformly distributed over $\{1,\dots,W_{mk}\}$. Each transmitter k consists of two[2] *encoders* f_{mk} of *block length* T, with

$$f_{mk} : \{1,\dots,W_{mk}\} \to Z_2^{n\times T},$$

mapping the message w_{mk} into the sequence of channel inputs

$$(\bar{\boldsymbol{u}}_{mk}[t])_{t=1}^T \triangleq f_{mk}(w_{mk}).$$

The rate of the encoder f_{mk} is

$$\bar{R}_{mk} \triangleq \frac{\log(W_{mk})}{T}.$$

A *decoder* of block length T at receiver m is a function

$$\phi_m : \mathbb{Z}_2^{n\times T} \to \{1,\dots,W_{m1}\} \times \{1,\dots,W_{m2}\}$$

mapping the sequence of channel outputs $(\bar{\boldsymbol{y}}_m[t])_{t=1}^T$ into the *message estimates*

$$(\hat{w}_{m1}, \hat{w}_{m2}) \triangleq \phi_m\big((\bar{\boldsymbol{y}}_m[t])_{t=1}^T\big)$$

of the messages (w_{m1}, w_{m2}). The *probability of error* of a pair of encoders (f_1, f_2) and decoders (ϕ_1, ϕ_2) is

$$\mathbb{P}\Big(\bigcup_{m,k}\{\hat{w}_{mk} \neq w_{mk}\}\Big).$$

Definition. A sum rate $\bar{R}(n)$ is *achievable*, if for every $\varepsilon > 0$ and every $T \geq T_0(\varepsilon)$ there exists encoders $(f_{11}, f_{12}, f_{21}, f_{22})$ with block length T, unit average power constraint, and rates $\sum_{m,k} \bar{R}_k \geq \bar{R}(n)$, and there exists decoders (ϕ_1, ϕ_2) with block length T and probability of error less than ε. The *sum capacity* $\bar{C}(n)$ of the (modulated) deterministic X-channel (15) is the supremum of all achievable sum rates $\bar{R}(n)$.

[1] This definition of the deterministic model corresponds to a power constraint of 16 in the Gaussian model. This is mainly for convenience of notation. Since the additional factor 16 in power only increases capacity by a constant number of bits per channel use, this does not significantly affect the quality of approximation.

[2] Observe that in the definition of capacity $\bar{C}(n)$ of the modulated deterministic X-channel (15) we use two encoders at each transmitter (one for the each of the two messages). This differs from the definition of capacity $C(n)$ of the Gaussian X-channel (9), where we use a single encoder. Thus, in the deterministic case, we force the messages to be encoded separately, while we allow joint encoding of the two messages in the Gaussian case. This restriction is introduced because the aim of the deterministic model is to better understand the modulated Gaussian X-channel (12), which already handles the joint encoding of the messages through the modulation process.

C. *X-Channel with Arbitrary* SNR*s*

In the last section, we introduced the Gaussian X-channel with SNR of order 2^{2n} over each of the four links. Thus, all links had approximately the same strength. We now introduce the Gaussian X-channel with arbitrary SNRs. The channel output y_m at receiver $m \in \{1, 2\}$ and time $t \in \mathbb{N}$ is then

$$y_m[t] \triangleq 2^{n_{m1}} h_{m1} x_1[t] + 2^{n_{m2}} h_{m2} x_2[t] + z_m[t], \tag{16}$$

where $x_k[t]$ is the channel input at transmitter $k \in \{1, 2\}$, where $2^{n_{mk}} h_{mk}$ is the channel gain from transmitter k to receiver m, and where $z_m[t] \sim \mathcal{N}(0, 1)$ is additive white Gaussian receiver noise. The channel gains consist again of two parts, $2^{n_{mk}}$ and h_{mk}. We assume that $n_{mk} \in \mathbb{Z}_+$ and that $h_{mk} \in (1, 2]$ for each m, k. Since $2^{n_{mk}} h_{mk}$ varies over $(2^{n_{mk}}, 2^{n_{mk}+1}]$ when h_{mk} varies over $(1, 2]$, we see that any real channel gain greater than one can be written in this form. As discussed in Section II-A, channel gains with magnitude less than one are not relevant for a constant-gap capacity approximation, and hence are ignored here. Similarly, negative channel gains have no effect on the achievable schemes and outer bounds presented later, and are therefore ignored as well. In other words, Equation (16) models essentially the general Gaussian X-channel. As before, the parameter n_{mk} captures the magnitude or coarse structure of the channel gain, the parameter h_{mk} captures the fine structure of the channel gain. Denote by

$$\boldsymbol{N} \triangleq \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix}$$

the collection of channel parameters.

Each transmitter has again one message to communicate to each receiver. So there are a total of four messages, denoted by w_{mk} for $m, k \in \{1, 2\}$. We impose a unit average power constraint on each transmitted signal. Define the sum capacity $C(\boldsymbol{N})$ of the Gaussian X-channel with arbitrary SNRs in analogy to the symmetric case.

We use the same modulation scheme (10) as in Section III-B to map the message w_{mk} into the real number x_{mk}. Transmitters one and two form the channel inputs

$$\begin{aligned} x_1 &\triangleq h_{22} u_{11} + h_{12} u_{21}, \\ x_2 &\triangleq h_{11} u_{22} + h_{21} u_{12}, \end{aligned}$$

The received signals are then given by

$$\begin{aligned} y_1 &= 2^{n_{11}} h_{11} h_{22} u_{11} + 2^{n_{12}} h_{12} h_{21} u_{12} + h_{11} h_{12} (2^{n_{11}} u_{21} + 2^{n_{12}} u_{22}) + z_1, \\ y_2 &= 2^{n_{22}} h_{22} h_{11} u_{22} + 2^{n_{21}} h_{12} h_{21} u_{21} + h_{22} h_{21} (2^{n_{22}} u_{12} + 2^{n_{21}} u_{11}) + z_2. \end{aligned}$$

Writing g_{mk} for the product of the channel gains as defined in (13), we can rewrite this as

$$\begin{aligned} y_1 &= 2^{n_{11}} g_{11} u_{11} + 2^{n_{12}} g_{12} u_{12} + g_{10} (2^{n_{11}} u_{21} + 2^{n_{12}} u_{22}) + z_1, \\ y_2 &= 2^{n_{22}} g_{22} u_{22} + 2^{n_{21}} g_{21} u_{21} + g_{20} (2^{n_{22}} u_{12} + 2^{n_{21}} u_{11}) + z_2. \end{aligned}$$

To gain intuition for the modulated Gaussian X-channel, we develop again the corresponding lower-triangular deterministic channel model. To simplify the presentation, we focus in the following on the case where the direct links are stronger than the cross links[3], i.e.,

$$\min\{n_{11}, n_{22}\} \geq \max\{n_{12}, n_{21}\}.$$

[3]This assumption is only made for ease of exposition. Since the labeling of the receivers is arbitrary, all results carry immediately over to the case $\min\{n_{12}, n_{21}\} \geq \max\{n_{11}, n_{22}\}$. The models and tools developed in this paper for these two cases can be applied to the other cases as well.

Following the derivation in Section III-B, we form the deterministic channel with inputs $\bar{\boldsymbol{u}}_{mk} \in \mathbb{Z}_2^{n_{kk}}$ for $m, k \in \{1, 2\}$ and with channel outputs $\bar{\boldsymbol{y}}_m \in \mathbb{Z}_2^{n_{mm}}$ for $m \in \{1, 2\}$. It will be convenient to split the channel input into "common" and "private" portions, i.e.,

$$\bar{\boldsymbol{u}}_{mk} \triangleq \begin{pmatrix} \bar{\boldsymbol{u}}_{mk}^{\mathsf{C}} \\ \bar{\boldsymbol{u}}_{mk}^{\mathsf{P}} \end{pmatrix},$$

where $\bar{\boldsymbol{u}}_{m1}^{\mathsf{C}} \in \mathbb{Z}_2^{n_{21}}$ and $\bar{\boldsymbol{u}}_{m2}^{\mathsf{C}} \in \mathbb{Z}_2^{n_{12}}$ for $m \in \{1, 2\}$. The deterministic channel connecting $\bar{\boldsymbol{u}}_{mk}$ with $\bar{\boldsymbol{y}}_m$ is then given by

$$\bar{\boldsymbol{y}}_1 \triangleq \bar{\boldsymbol{G}}_{11}\bar{\boldsymbol{u}}_{11} \oplus \bar{\boldsymbol{G}}_{12} \begin{pmatrix} \boldsymbol{0} \\ \bar{\boldsymbol{u}}_{12}^{\mathsf{C}} \end{pmatrix} \oplus \bar{\boldsymbol{G}}_{10} \left(\bar{\boldsymbol{u}}_{21} \oplus \begin{pmatrix} \boldsymbol{0} \\ \bar{\boldsymbol{u}}_{22}^{\mathsf{C}} \end{pmatrix} \right), \tag{17a}$$

$$\bar{\boldsymbol{y}}_2 \triangleq \bar{\boldsymbol{G}}_{22}\bar{\boldsymbol{u}}_{22} \oplus \bar{\boldsymbol{G}}_{21} \begin{pmatrix} \boldsymbol{0} \\ \bar{\boldsymbol{u}}_{21}^{\mathsf{C}} \end{pmatrix} \oplus \bar{\boldsymbol{G}}_{20} \left(\bar{\boldsymbol{u}}_{12} \oplus \begin{pmatrix} \boldsymbol{0} \\ \bar{\boldsymbol{u}}_{11}^{\mathsf{C}} \end{pmatrix} \right), \tag{17b}$$

see Figs. 5 and 6. Here, the lower-triangular binary matrices $\bar{\boldsymbol{G}}_{mk}$ are defined in analogy to (14). The matrix $\bar{\boldsymbol{G}}_{1k}$ is of dimension $n_{11} \times n_{11}$ and $\bar{\boldsymbol{G}}_{2k}$ is of dimension $n_{22} \times n_{22}$ for all $k \in \{0, 1, 2\}$. As in the symmetric case, we assume for the deterministic model that $g_{mk} \in (1, 2]$. As a consequence, the diagonal elements of $\bar{\boldsymbol{G}}_{mk}$ are equal to one.

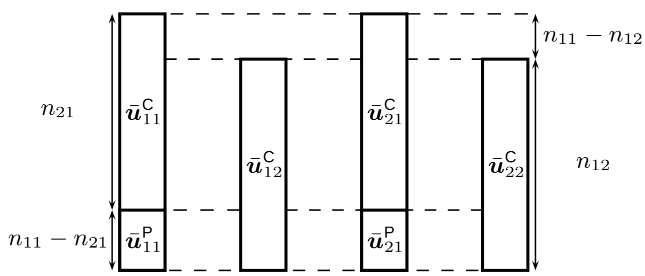

Fig. 5. Deterministic model at receiver one. For simplicity, the matrices $\bar{\boldsymbol{G}}_{mk}$ are omitted. The interference terms $\bar{\boldsymbol{u}}_{21}$ and $\bar{\boldsymbol{u}}_{22}$ are observed at receiver one multiplied by the same matrix $\bar{\boldsymbol{G}}_{10}$. The desired terms $\bar{\boldsymbol{u}}_{11}$ and $\bar{\boldsymbol{u}}_{12}$ are multiplied by different matrices $\bar{\boldsymbol{G}}_{11}$ and $\bar{\boldsymbol{G}}_{12}$, respectively.

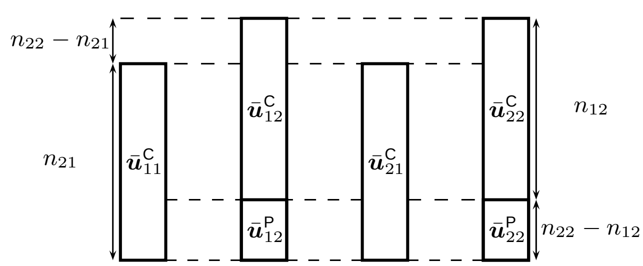

Fig. 6. Deterministic model at receiver two. The matrices $\bar{\boldsymbol{G}}_{mk}$ are again omitted. The interference terms $\bar{\boldsymbol{u}}_{11}$ and $\bar{\boldsymbol{u}}_{12}$ are observed at receiver one multiplied by the same matrix $\bar{\boldsymbol{G}}_{20}$. The desired terms $\bar{\boldsymbol{u}}_{21}$ and $\bar{\boldsymbol{u}}_{22}$ are multiplied by different matrices $\bar{\boldsymbol{G}}_{21}$ and $\bar{\boldsymbol{G}}_{22}$, respectively.

The definition of an achievable sum rate $\bar{R}(\boldsymbol{N})$ and of capacity $\bar{C}(\boldsymbol{N})$ for the deterministic X-channel with arbitrary SNRs is in analogy to the symmetric case.

IV. Main Results

The main result of this paper is a constant-gap approximation for the capacity of the Gaussian X-channel. To simplify the presentation of the relevant concepts and results, we start with the analysis of the Gaussian X-channel with symmetric SNRs in Section IV-A. We then consider the Gaussian X-channel with arbitrary SNRs in Section IV-B.

A. X-Channel with Symmetric SNRs

We start with the analysis of the deterministic X-channel—as we will see in the following, the insights obtained for this model carry over to the Gaussian X-channel. The capacity $\bar{C}(n)$ of the symmetric deterministic X-channel is characterized by the next theorem.

Theorem 1. *For every $\delta \in (0,1]$ and $n \in \mathbb{Z}_+$, there exists a set $B_n \subseteq (1,2]^{2\times 3}$ of Lebesgue measure*

$$\mu(B_n) \leq \delta$$

such that for all channel gains $(g_{mk}) \in (1,2]^{2\times 3} \setminus B_n$ the sum capacity $\bar{C}(n)$ of the (modulated) symmetric deterministic X-channel (15) *satisfies*

$$\tfrac{4}{3}n - 2\log(c_1/\delta) \leq \bar{C}(n) \leq \tfrac{4}{3}n$$

for some finite positive universal constant c_1.

Theorem 1 is a special case of Theorem 5 presented in Section IV-B. We hence omit the proof of Theorem 1.

Theorem 1 approximates the capacity of the modulated deterministic X-channel (15) up to a constant gap for all channel gains $g_{mk} \in (1,2]$ except for a set B_n of arbitrarily small measure. As will become clear in the following, the event $(g_{mk}) \in B_n$ can be interpreted as an outage event, as in this case the proposed achievable scheme fails to deliver the target rate of $\frac{4}{3}n - 2\log(c_1/\delta)$. Here δ parametrizes the trade-off between the measure of the outage event and the target rate: decreasing δ decreases the measure of the outage event B_n, but at the same time also decreases the target rate $\frac{4}{3}n - 2\log(c_1/\delta)$. We point out that δ can be chosen independently of the number of input bits n, hence the approximation gap is uniform in n.

Theorem 1 can be used to derive the more familiar result on the degrees-of-freedom of the deterministic X-channel. Setting $\delta = n^{-2}$ results in the measures $\mu(B_n) \leq n^{-2}$ to be summable over $n \in \mathbb{Z}_+$. Applying the Borel-Cantelli lemma yields then the following corollary to Theorem 1.

Corollary 2. *For almost all channel gains $(g_{mk}) \in (1,2]^{2\times 3}$ the sum capacity $\bar{C}(n)$ of the (modulated) symmetric deterministic X-channel* (15) *satisfies*

$$\lim_{n\to\infty} \frac{\bar{C}(n)}{n} = 4/3.$$

Corollary 2 states that for almost all channel gains the deterministic X-channel has $4/3$ degrees-of-freedom. We emphasize that, while Corollary 2 is simpler to state and perhaps more familiar in form, Theorem 1 is considerably stronger. Indeed, Theorem 1 provides the stronger *constant gap* capacity approximation for the sum capacity $\bar{C}(n)$, whereas Corollary 2 provides the weaker *degrees-of-freedom* capacity approximation. Moreover, Theorem 1 provides bounds for *finite* n on the measure of the outage event B_n, whereas Corollary 2 provides only *asymptotic* information about its size.

We now describe the communication scheme achieving the lower bound in Theorem 1 (see Fig. 7). Use the first $\bar{R}$ components of each vector $\bar{\boldsymbol{u}}_{mk}$ to transmit information, and set the last $n - \bar{R}$ components to zero. The rate of this communication scheme is hence $4\bar{R}$. Receiver one is interested in $\bar{\boldsymbol{u}}_{11}$ and $\bar{\boldsymbol{u}}_{12}$. These vectors are received in the subspace spanned by the first $\bar{R}$ columns of $\bar{\boldsymbol{G}}_{11}$ and $\bar{\boldsymbol{G}}_{12}$, respectively. On the other hand, the messages $\bar{\boldsymbol{u}}_{21}$ and $\bar{\boldsymbol{u}}_{22}$ that receiver one is not interested in, and that can hence be regarded as interference, are both received in the same subspace spanned by the first $\bar{R}$ columns of $\bar{\boldsymbol{G}}_{10}$. Thus, the two interference vectors are aligned in a subspace of dimension $\bar{R}$. The situation at receiver two is similar.

Assume at receiver one the three subspaces spanned by the first $\bar{R}$ columns of $\bar{\boldsymbol{G}}_{11}$, $\bar{\boldsymbol{G}}_{12}$, and $\bar{\boldsymbol{G}}_{10}$ were linearly independent. Then the receiver could recover the two desired vectors by projecting the received vector into the corresponding subspaces in order to zero force the two interfering vectors. We show that for most channel gains this linear independence of the three subspaces holds for $\bar{R} \approx n/3$. The outage

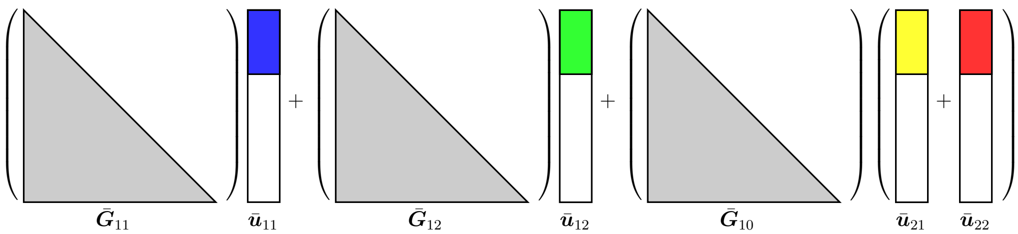

Fig. 7. Allocation of bits for the deterministic X-channel with symmetric SNRs as seen at receiver one. The white regions correspond to zero bits; the shaded regions carry information.

event B_n is thus precisely the event that at either of the two receivers the three subspaces spanned by the first $\bar{R}$ columns of $\bar{\boldsymbol{G}}_{m1}$, $\bar{\boldsymbol{G}}_{m2}$, and $\bar{\boldsymbol{G}}_{m0}$ are not linearly independent.

The results for the deterministic X-channel suggest that the modulation scheme (11) should achieve a sum rate of

$$3\bar{R} \approx \tfrac{4}{3}n \pm O(1)$$

as $n \to \infty$. Furthermore, it suggests that a n-bit quantization of the channel gains h_{mk} available at both transmitters and receivers should be sufficient to achieve this asymptotic rate. This intuition turns out to be correct, as the next theorem shows.

Theorem 3. *For every $\delta \in (0,1]$ and $n \in \mathbb{Z}_+$, there exists a set $B_n \subseteq (1,2]^{2\times 2}$ of Lebesgue measure at most*

$$\mu(B_n) \leq \delta$$

such that for all channel gains $(h_{mk}) \in (1,2]^{2\times 2} \setminus B_n$ the sum capacity of the symmetric Gaussian X-channel (9) *satisfies*

$$\tfrac{4}{3}n - 2\log(c_2/\delta) \leq C(n) \leq \tfrac{4}{3}n + 4$$

for some finite positive universal constant c_2. Moreover, the lower bound is achievable with a n-bit quantization of the channel gains h_{mk} available at both transmitters and receivers.

Theorem 3 is a special case of Theorem 6 presented in Section IV-B. We hence omit the proof of Theorem 3.

Theorem 3 provides a constant-gap capacity approximation for the symmetric Gaussian X-channel. The constant in the approximation is uniform in the channel gains $h_{mk} \in (1,2]$ except for a set B_n of arbitrarily small measure, and uniform in the power constraint n. The event B_n can again be interpreted as an outage event, and δ parametrizes the trade-off between the measure of the outage event B_n and the target rate of the achievable scheme. Since δ can be chosen independently of n, the approximation gap is uniform in the SNR, i.e., uniform in 2^{2n}.

Remark: It is worth pointing out that the set B_n can be explicitly computed: Given channel gains h_{mk}, there is an algorithm that can determine in bounded time if these channel gains are in the outage set B_n. More precisely, B_n is the union of $2^{\Theta(n)}$ diagonal strips (see the proof of Lemma 14 in Section VII-B and Fig. 17 there). Membership of (h_{mk}) in the outage set B_n is mostly determined by the $\Theta(n)$ most-significant bits in the binary expansion of the channel gains h_{mk}. In particular, for any finite n (and hence finite SNR), the question of rationality or irrationality of the channel gains h_{mk} is largely irrelevant to determining membership in B_n.

The theorem shows furthermore that the proposed achievable scheme for the Gaussian X-channel is not dependent on the exact knowledge of the channel gains, and a quantized version, available at all transmitters and receivers, is sufficient. In fact, the scheme achieving the lower bound uses mismatched encoders and decoders. The encoders perform modulation with respect to the *wrong* channel model

$$y_m[t] = 2^n\hat{h}_{m1}x_1[t] + 2^n\hat{h}_{m2}x_2[t] + z_m[t], \tag{18}$$

where $\hat{h}_{mk}$ is a n-bit (or, equivalently, $\frac{1}{2}\log(\mathsf{SNR})$-bit) quantization of the true channel gain h_{mk}. In other words, the channel inputs are

$$x_1[t] = \hat{h}_{22}u_{11}[t] + \hat{h}_{12}u_{21}[t],$$
$$x_2[t] = \hat{h}_{11}u_{22}[t] + \hat{h}_{21}u_{12}[t].$$

The decoders perform maximum-likelihood decoding also with respect to the wrong channel model (18). Thus, both the encoders and the decoders treat the channel estimates as if they were the true channel gains. This shows that the proposed achievable scheme is actually quite robust with respect to channel estimation and quantization errors.

As before, we can use Theorem 3 to derive more familiar results on the degrees-of-freedom of the Gaussian X-channel. Consider a sequence of SNRs 2^{2n} indexed by $n \in \mathbb{Z}_+$, and set $\delta = n^{-2}$. Then the measures $\mu(B_n) \leq n^{-2}$ are summable over $n \in \mathbb{Z}_+$. Applying the Borel-Cantelli lemma as before yields the following corollary to Theorem 3.

Corollary 4. *For almost all channel gains* $(h_{mk}) \in (1,2]^{2\times 2}$ *the sum capacity* $C(n)$ *of the symmetric Gaussian X-channel* (9) *satisfies*

$$\lim_{n\to\infty} \frac{C(n)}{n} = 4/3.$$

Since the SNR of the channel is approximately 2^{2n} so that $n \approx \frac{1}{2}\log(\mathsf{SNR})$, Corollary 4 shows that the Gaussian X-channel has $4/3$ degrees-of-freedom for almost all channel gains h_{mk}, recovering the result in [3]. We emphasize again that Theorem 3 is considerably stronger than Corollary 4. Indeed, Theorem 3 proves the *constant-gap* capacity approximation

$$|C(n) - \tfrac{4}{3}n| \leq O(1)$$

with pre-constant in the $O(1)$ term uniform in the channel gains h_{mk} outside B_n. This is considerably stronger than the *degrees-of-freedom* capacity approximation in Corollary 4, which shows only that

$$|C(n) - \tfrac{4}{3}n| \leq o(n)$$

with pre-constant in the $o(n)$ term depending on h_{mk}. Moreover, Theorem 3 provides bounds on the measure of the outage event for *finite* SNRs, not just *asymptotic* guarantees as in Corollary 4.

B. X-Channel with Arbitrary SNR*s*

In the last section, we considered the Gaussian X-channel with SNRs across each link of order 2^{2n}. Thus, all links had approximately the same strength. We now turn to the Gaussian X-channel with arbitrary SNRs. As before, we start with the analysis of the deterministic X-channel. The next theorem provides an approximate characterization of the sum capacity $\bar{C}(\boldsymbol{N})$ of the general deterministic X-channel with bit levels $\boldsymbol{N}$.

Theorem 5. *For every* $\delta \in (0,1]$ *and* $\boldsymbol{N} \in \mathbb{Z}_+^{2\times 2}$ *such that* $\min\{n_{11}, n_{22}\} \geq \max\{n_{12}, n_{21}\}$ *there exists a set* $B \subseteq (1,2]^{2\times 3}$ *of Lebesgue measure*

$$\mu(B) \leq \delta$$

such that for all channel gains $(g_{mk}) \in (1,2]^{2\times 3} \setminus B$ *the sum capacity* $\bar{C}(\boldsymbol{N})$ *of the (modulated) general deterministic X-channel* (17) *satisfies*

$$\bar{D}(\boldsymbol{N}) - 2\log(c_1/\delta) \leq \bar{C}(\boldsymbol{N}) \leq \bar{D}(\boldsymbol{N})$$

for some finite positive universal constant c_1*, and where*

$$\bar{D}(\boldsymbol{N}) \triangleq \min\big\{\bar{D}_1(\boldsymbol{N}), \bar{D}_2(\boldsymbol{N}), \bar{D}_3(\boldsymbol{N}), \bar{D}_4(\boldsymbol{N})\big\} + (n_{11} - n_{21}) + (n_{22} - n_{12})$$

and

$$\bar{D}_1(\boldsymbol{N}) \triangleq (n_{12}+n_{21}-n_{11})^+ + (n_{12}+n_{21}-n_{22})^+,$$
$$\bar{D}_2(\boldsymbol{N}) \triangleq \tfrac{1}{2}\big(n_{12}+n_{21}+(n_{12}+n_{21}-n_{22})^+\big),$$
$$\bar{D}_3(\boldsymbol{N}) \triangleq \tfrac{1}{2}\big(n_{12}+n_{21}+(n_{12}+n_{21}-n_{11})^+\big),$$
$$\bar{D}_4(\boldsymbol{N}) \triangleq \tfrac{2}{3}(n_{12}+n_{21}).$$

The proof of Theorem 5 is presented in Section V. For the special case of symmetric channel SNRs, $n_{mk}=n$ for all m,k, Theorem 5 reduces to Theorem 1 in Section IV-A.

We now provide a sketch of the communication scheme achieving the lower bound in Theorem 5 (see Figs. 8 and 9). Observe that the $n_{11}-n_{21}$ least significant bits $\bar{\boldsymbol{u}}_{11}^{\mathsf{P}}$ of $\bar{\boldsymbol{u}}_{11}$ are not visible at the second receiver. Therefore, we can use these bits to privately carry $n_{11}-n_{21}$ bits from the first transmitter to the first receiver without affecting the second receiver. The rate of this private message is denoted by $\bar{R}_{11}^{\mathsf{P}}$. The remaining rate is denoted by $\bar{R}_{11}^{\mathsf{C}}$, i.e.,

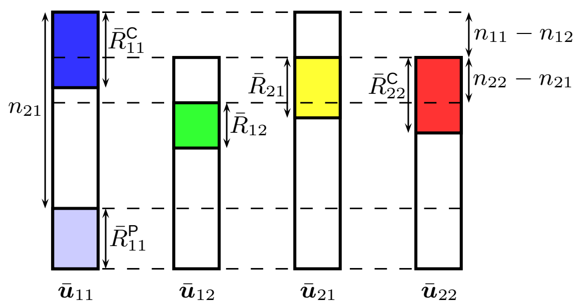

Fig. 8. Allocation of bits at receiver one. Here, $\bar{\boldsymbol{u}}_{11}$ and $\bar{\boldsymbol{u}}_{12}$ are the desired bits and are received multiplied by the matrices $\bar{\boldsymbol{G}}_{11}$ and $\bar{\boldsymbol{G}}_{12}$ (not shown in the figure), respectively. The vectors $\bar{\boldsymbol{u}}_{21}$ and $\bar{\boldsymbol{u}}_{22}$ are interference and are both received multiplied by the same matrix $\bar{\boldsymbol{G}}_{10}$.

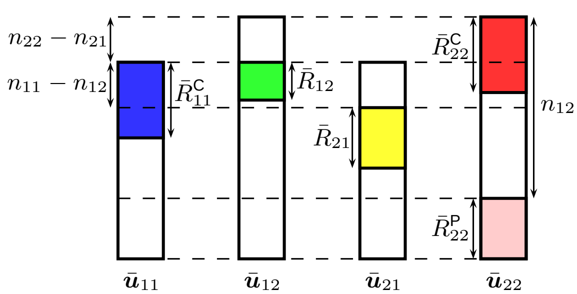

Fig. 9. Allocation of bits at receiver two. Here, $\bar{\boldsymbol{u}}_{21}$ and $\bar{\boldsymbol{u}}_{22}$ are the desired bits and are received multiplied by the matrices $\bar{\boldsymbol{G}}_{21}$ and $\bar{\boldsymbol{G}}_{22}$ (not shown in the figure), respectively. The vectors $\bar{\boldsymbol{u}}_{11}$ and $\bar{\boldsymbol{u}}_{12}$ are interference and are both received multiplied by the same matrix $\bar{\boldsymbol{G}}_{20}$.

$$\bar{R}_{11} \triangleq \bar{R}_{11}^{\mathsf{C}} + \bar{R}_{11}^{\mathsf{P}},$$

where

$$\bar{R}_{11}^{\mathsf{P}} \triangleq n_{11}-n_{21}.$$

Similarly, the $n_{22}-n_{12}$ least significant bits $\bar{\boldsymbol{u}}_{22}^{\mathsf{P}}$ of $\bar{\boldsymbol{u}}_{22}$ are not visible at the second receiver. Therefore, we can use this part to privately carry $n_{22}-n_{12}$ bits from the second transmitter to the second receiver without affecting the first receiver. The rate of this private message is denoted by $\bar{R}_{22}^{\mathsf{P}}$. The remaining rate is denoted by $\bar{R}_{22}^{\mathsf{C}}$, i.e.,

$$\bar{R}_{22} \triangleq \bar{R}_{22}^{\mathsf{C}} + \bar{R}_{22}^{\mathsf{P}},$$

where

$$\bar{R}_{22}^{\mathsf{P}} \triangleq n_{22} - n_{12}.$$

It remains to choose the values of $\bar{R}_{11}^{\mathsf{C}}$, $\bar{R}_{22}^{\mathsf{C}}$, $\bar{R}_{12}$, and $\bar{R}_{21}$. Our proposed design rules are as follows.

- We dedicate the $\bar{R}_{11}^{\mathsf{C}}$ most significant bits of $\bar{\boldsymbol{u}}_{11}$ to carry information from transmitter one to receiver one.
- Similarly, we dedicate the $\bar{R}_{22}^{\mathsf{C}}$ most significant bits of $\bar{\boldsymbol{u}}_{22}$ to carry information from transmitter two to receiver two.
- We always set the $n_{22} - n_{21}$ most significant bits of $\bar{\boldsymbol{u}}_{12}$ to zero. The next $\bar{R}_{12}$ bits of $\bar{\boldsymbol{u}}_{12}$ carry information from transmitter two to receiver one. As shown in Fig. 9, this guarantees the (partial) alignment of $\bar{\boldsymbol{u}}_{12}$ with $\bar{\boldsymbol{u}}_{11}$ at the second receiver.
- We always set the $n_{11} - n_{12}$ most significant bits of $\bar{\boldsymbol{u}}_{21}$ to zero. The next $\bar{R}_{21}$ bits of $\bar{\boldsymbol{u}}_{21}$ carry information from transmitter one to receiver two. As shown in Fig. 8, this guarantees the (partial) alignment of $\bar{\boldsymbol{u}}_{21}$ with $\bar{\boldsymbol{u}}_{22}$ at the first receiver.

Optimizing the values of the rates $\bar{R}_{mk}$ subject to the condition that both receivers can decode the desired messages, yields the lower bound in Theorem 5. The details of this analysis can be found in Section V-A.

Generalizing these ideas from the deterministic to the Gaussian model, we obtain the following constant-gap capacity approximation for the Gaussian X-channel with general asymmetric channel gains.

Theorem 6. *For every $\delta \in (0,1]$ and $\boldsymbol{N} \in \mathbb{Z}_+^{2\times 2}$ such that $\min\{n_{11}, n_{22}\} \geq \max\{n_{12}, n_{21}\}$ there exists a set $B \subseteq (1,2]^{2\times 2}$ of Lebesgue measure*

$$\mu(B) \leq \delta$$

such that for all channel gains $(h_{mk}) \in (1,2]^{2\times 2} \setminus B$ the sum capacity $C(\boldsymbol{N})$ of the general Gaussian X-channel (16) *satisfies*

$$D(\boldsymbol{N}) - 2\log(c_2/\delta) \leq C(\boldsymbol{N}) \leq D(\boldsymbol{N}) + 4$$

for some finite positive universal constant c_2, and where

$$D(\boldsymbol{N}) \triangleq \min\big\{D_1(\boldsymbol{N}), D_2(\boldsymbol{N}), D_3(\boldsymbol{N}), D_4(\boldsymbol{N})\big\} + (n_{11} - n_{21}) + (n_{22} - n_{12})$$

and

$$\begin{aligned} D_1(\boldsymbol{N}) &\triangleq (n_{12} + n_{21} - n_{11})^+ + (n_{12} + n_{21} - n_{22})^+, \\ D_2(\boldsymbol{N}) &\triangleq \tfrac{1}{2}\big(n_{12} + n_{21} + (n_{12} + n_{21} - n_{22})^+\big), \\ D_3(\boldsymbol{N}) &\triangleq \tfrac{1}{2}\big(n_{12} + n_{21} + (n_{12} + n_{21} - n_{11})^+\big), \\ D_4(\boldsymbol{N}) &\triangleq \tfrac{2}{3}(n_{12} + n_{21}). \end{aligned}$$

Moreover, the lower bound on $C(\boldsymbol{N})$ is achievable with a $\max\{n_{mk}\}$-bit quantization of the channel gains h_{mk} available at both transmitters and receivers.

The proof of Theorem 6 is presented in Section VI. For the special case of symmetric channel SNRs, $n_{mk} = n$ for all m, k, Theorem 6 reduces to Theorem 3 in Section IV-A. We point out that $D(\boldsymbol{N})$ in Theorem 6 for the Gaussian X-channel is equal to $\bar{D}(\boldsymbol{N})$ in Theorem 5 for the deterministic X-channel. Thus, up to a constant gap, the Gaussian X-channel and its lower-triangular deterministic approximation have the same capacity.

The lower bound in Theorem 6 is achieved by encoders and decoders that have access to only a $\max\{n_{mk}\}$-bit quantization $\hat{h}_{mk}$ of the channel gains h_{mk}. As before, the encoders and decoders are mismatched, in the sense that they are operating under the assumption that $\hat{h}_{mk}$ is the correct channel gain. This shows again that the proposed communication scheme is quite robust with respect to channel estimation and quantization errors.

V. Proof of Theorem 5 (Deterministic X-Channel)

This section contains the proof of the capacity approximation for the deterministic X-channel in Theorem 5. Achievability of the lower bound in the theorem is proved in Section V-A, the upper bound is proved in Section V-B.

A. Achievability for the Deterministic X-Channel

This section contains the proof of the lower bound in Theorem 5. Without loss of generality, we assume that $n_{22} \geq n_{11}$. We use the achievable scheme outlined in Section IV-B (see Figs. 8 and 9 there). We want to maximize the sum rate

$$\bar{R}_{11}^{\mathsf{C}} + \bar{R}_{11}^{\mathsf{P}} + \bar{R}_{22}^{\mathsf{C}} + \bar{R}_{22}^{\mathsf{P}} + \bar{R}_{12} + \bar{R}_{21},$$

where

$$\bar{R}_{kk}^{\mathsf{C}} + \bar{R}_{kk}^{\mathsf{P}} = \bar{R}_{kk}$$

is the total rate from transmitter k to receiver k. The constraint is that each receiver can solve for its own desired messages plus the visible parts of the aligned interference bits.

If the subspaces spanned by the columns of $\bar{\boldsymbol{G}}_{mk}$ corresponding to information-bearing bits of $\bar{\boldsymbol{u}}_{mk}$ are linearly independent, then there exists a unique channel input to the deterministic X-channel that results in the observed channel output. The decoder declares that this unique channel input was sent. The next lemma provides a sufficient condition for this linear independence to hold and hence for decoding to be successful.

Lemma 7. *Let $\delta \in (0,1]$ and $\boldsymbol{N} \in \mathbb{Z}_+^{2\times 2}$ such that $\min\{n_{11}, n_{22}\} \geq \max\{n_{12}, n_{21}\}$. Assume $\bar{R}_{11}^{\mathsf{P}}$,$\bar{R}_{11}^{\mathsf{C}}$, $\bar{R}_{12}$, $\bar{R}_{21}$, $\bar{R}_{22}^{\mathsf{P}}$, $\bar{R}_{22}^{\mathsf{C}} \in \mathbb{Z}_+$ satisfy*

$$\bar{R}_{11}^{\mathsf{C}} + \max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11} - \log(32/\delta), \tag{19a}$$

$$\mathbb{1}_{\{\max\{\bar{R}_{12}, \bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\}>0\}} \cdot \big(\max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}}\big) \leq n_{12} - \log(32/\delta), \tag{19b}$$

$$\mathbb{1}_{\{\bar{R}_{12}>0\}} \cdot \big(\bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}}\big) \leq n_{12} + n_{21} - n_{22}, \tag{19c}$$

and

$$\bar{R}_{22}^{\mathsf{C}} + \max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}} \leq n_{22} - \log(32/\delta), \tag{20a}$$

$$\mathbb{1}_{\{\max\{\bar{R}_{21}, \bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\}>0\}} \cdot \big(\max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}}\big) \leq n_{21} - \log(32/\delta), \tag{20b}$$

$$\mathbb{1}_{\{\bar{R}_{21}>0\}} \cdot \big(\bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}}\big) \leq n_{12} + n_{21} - n_{11}. \tag{20c}$$

Then the bit allocation in Section IV-B for the (modulated) deterministic X-channel (17) *allows successful decoding at both receivers for all channel gains $(g_{mk}) \in (1,2]^{2\times 3}$ except for a set $B \subset (1,2]^{2\times 3}$ of Lebesgue measure*

$$\mu(B) \leq \delta.$$

The proof of Lemma 7 can be found in Section VII-A.

We now choose rates satisfying these decoding conditions. For ease of notation, we will ignore the $\log(32/\delta)$ terms throughout—the reduction in sum rate due to this additional requirement is at most $2\log(32/\delta)$. The optimal allocation of bits at the transmitters depends on the value $n_{12} + n_{21}$. We treat the cases

$$\begin{aligned}
\text{I:} &\quad n_{12} + n_{21} \in \big[0, n_{11}\big] \\
\text{II:} &\quad n_{12} + n_{21} \in \big(n_{11}, n_{22}\big] \\
\text{III:} &\quad n_{12} + n_{21} \in \big(n_{22}, n_{11} + \tfrac{1}{2}n_{22}\big] \\
\text{IV:} &\quad n_{12} + n_{21} \in \big(n_{11} + \tfrac{1}{2}n_{22}, \tfrac{3}{2}n_{22}\big] \\
\text{V:} &\quad n_{12} + n_{21} \in \big(\tfrac{3}{2}n_{22}, n_{11} + n_{22}\big]
\end{aligned}$$

separately. Since $n_{12}+n_{21} \leq n_{11}+n_{22}$ by the assumption $\max\{n_{12}, n_{21}\} \leq \min\{n_{11}, n_{22}\}$, this covers all possible values of $\boldsymbol{N}$.

Case I ($n_{12}+n_{21} \leq n_{11}$): We set

$$\begin{aligned}
\bar{R}^{\mathsf{P}}_{11} &\triangleq n_{11}-n_{21}, \\
\bar{R}^{\mathsf{P}}_{22} &\triangleq n_{22}-n_{12}, \\
\bar{R}^{\mathsf{C}}_{22} &\triangleq \bar{R}^{\mathsf{C}}_{11} \triangleq \bar{R}_{12} \triangleq \bar{R}_{21} \triangleq 0.
\end{aligned}$$

In words, we solely communicate using the private channel inputs $\bar{\boldsymbol{u}}^{\mathsf{P}}_{11}$ and $\bar{\boldsymbol{u}}^{\mathsf{P}}_{22}$. Recall that, by our assumptions throughout this section, $\max\{n_{12}, n_{21}\} \leq n_{11} \leq n_{22}$. Hence, $\bar{R}^{\mathsf{P}}_{11} \geq 0$ and $\bar{R}^{\mathsf{P}}_{22} \geq 0$, and this rate allocation is valid. The calculation in Appendix A verifies that this rate allocation satisfies the decoding conditions (19) and (20) in Lemma 7. Hence both receivers can recover the desired messages. The sum rate is

$$\begin{aligned}
(n_{11}-n_{21})+(n_{22}-n_{12}) &= (n_{12}+n_{21}-n_{11})^{+}+(n_{12}+n_{21}-n_{22})^{+}+(n_{11}-n_{21})+(n_{22}-n_{12}) \\
&= \bar{D}_1(\boldsymbol{N})+(n_{11}-n_{21})+(n_{22}-n_{12}) \\
&\geq \bar{D}(\boldsymbol{N}), \qquad (21)
\end{aligned}$$

where we have used that

$$n_{12}+n_{21} \leq n_{11} \leq n_{22}.$$

Case II ($n_{11} < n_{12}+n_{21} \leq n_{22}$): We set

$$\begin{aligned}
\bar{R}^{\mathsf{P}}_{11} &\triangleq n_{11}-n_{21}, \\
\bar{R}^{\mathsf{P}}_{22} &\triangleq n_{22}-n_{12}, \\
\bar{R}^{\mathsf{C}}_{22} &\triangleq n_{12}-\bar{R}^{\mathsf{P}}_{11}, \\
\bar{R}^{\mathsf{C}}_{11} &\triangleq \bar{R}_{12} \triangleq \bar{R}_{21} \triangleq 0,
\end{aligned}$$

as shown in Fig. 10. Since $n_{12}+n_{21} > n_{11}$, we have $\bar{R}^{\mathsf{C}}_{22} \geq 0$, and hence this rate allocation is valid.

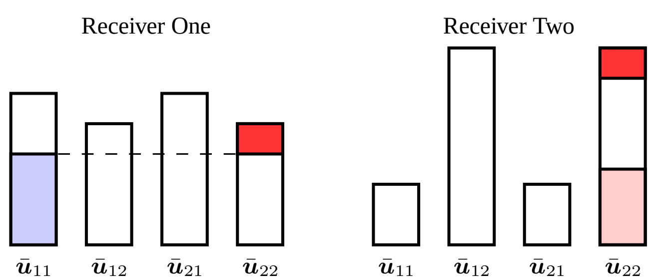

Fig. 10. Allocation of bits in case II. Here $n_{11}=10, n_{22}=13, n_{12}=8, n_{21}=4$. The transmitters send private messages at rates $\bar{R}^{\mathsf{P}}_{11}=6$ and $\bar{R}^{\mathsf{P}}_{22}=5$. Transmitter two sends a common message to receiver two at rate $\bar{R}^{\mathsf{C}}_{22}=2$.

The calculation in Appendix A verifies that this rate allocation satisfies the decoding conditions (19) and (20) in Lemma 7. Hence both receivers can decode successfully. The sum rate is

$$\begin{aligned}
&(n_{12}+n_{21}-n_{11})+(n_{11}-n_{21})+(n_{22}-n_{12}) \\
&\quad = (n_{12}+n_{21}-n_{11})^{+}+(n_{12}+n_{21}-n_{22})^{+}+(n_{11}-n_{21})+(n_{22}-n_{12}) \\
&\quad = \bar{D}_1(\boldsymbol{N})+(n_{11}-n_{21})+(n_{22}-n_{12}) \\
&\quad \geq \bar{D}(\boldsymbol{N}), \qquad (22)
\end{aligned}$$

where we have used that

$$n_{12}+n_{21} \leq n_{22}.$$

Case III ($n_{22} < n_{12} + n_{21} \le n_{11} + \frac{1}{2}n_{22}$): We set

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11} - n_{21}, \\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22} - n_{12}, \\
\bar{R}_{12} &\triangleq (n_{12} + 2n_{21} - n_{11} - n_{22})^+, \\
\bar{R}_{21} &\triangleq (n_{21} + 2n_{12} - n_{11} - n_{22})^+, \\
\bar{R}_{11}^{\mathsf{C}} &\triangleq n_{21} - \bar{R}_{22}^{\mathsf{P}} - \bar{R}_{21}, \\
\bar{R}_{22}^{\mathsf{C}} &\triangleq n_{12} - \bar{R}_{11}^{\mathsf{P}} - \bar{R}_{12},
\end{aligned}$$

as depicted in Fig. 11. Using $n_{12} + n_{21} > n_{22}$ and $n_{22} \ge n_{11} \ge \max\{n_{12}, n_{21}\}$, it can be verified that $\bar{R}_{11}^{\mathsf{C}} \ge 0$ and $\bar{R}_{22}^{\mathsf{C}} \ge 0$, and hence this rate allocation is valid.

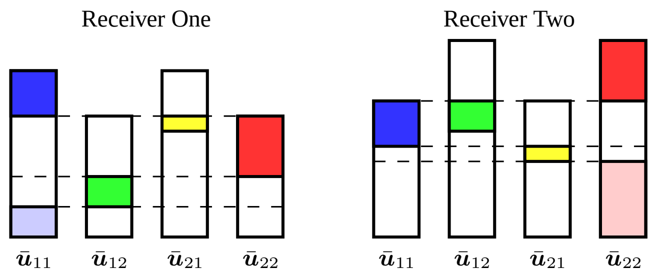

Fig. 11. Allocation of bits in case III. Here $n_{11} = 11, n_{22} = 13, n_{12} = 8, n_{21} = 9$. The transmitters send private messages at rates $\bar{R}_{11}^{\mathsf{P}} = 2$ and $\bar{R}_{22}^{\mathsf{P}} = 5$. Transmitter one sends a common message to receiver one at rate $\bar{R}_{11}^{\mathsf{C}} = 3$. Transmitter two sends a common message to receiver two at rate $\bar{R}_{22}^{\mathsf{C}} = 4$. The rates over the cross links are $\bar{R}_{12} = 2$ and $\bar{R}_{21} = 1$. Observe that the interference terms are partially aligned at each receiver.

The calculation in Appendix A verifies the decoding conditions (19) and (20) in Lemma 7. The sum rate of this allocation is

$$\begin{aligned}
n_{12} + n_{21} &= (n_{12} + n_{21} - n_{22}) + (n_{12} + n_{21} - n_{11}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\
&= (n_{12} + n_{21} - n_{11})^+ + (n_{12} + n_{21} - n_{22})^+ + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\
&= \bar{D}_1(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\
&\ge \bar{D}(\boldsymbol{N}),
\end{aligned} \tag{23}$$

where we have used that

$$n_{12} + n_{21} \ge n_{22} \ge n_{11}.$$

Case IV ($n_{11} + \frac{1}{2}n_{22} < n_{12} + n_{21} \le \frac{3}{2}n_{22}$): We set

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11} - n_{21}, \\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22} - n_{12}, \\
\bar{R}_{21} &\triangleq \left\lfloor n_{12} - \tfrac{1}{2}n_{22} \right\rfloor, \\
\bar{R}_{12} &\triangleq \bar{R}_{11}^{\mathsf{C}} \triangleq \left\lfloor n_{21} - \tfrac{1}{2}n_{22} \right\rfloor, \\
\bar{R}_{22}^{\mathsf{C}} &\triangleq n_{22} - n_{21},
\end{aligned}$$

as shown in Fig. 12. Note that

$$n_{11} + \tfrac{1}{2}n_{22} \le n_{12} + n_{21} \le n_{12} + n_{11}$$

and

$$n_{11} + \tfrac{1}{2}n_{22} \le n_{12} + n_{21} \le n_{11} + n_{21}$$

Therefore,

$$\tfrac{1}{2}n_{22} \leq \min\{n_{12}, n_{21}\}. \tag{24}$$

It follows that $\bar{R}_{21}$, $\bar{R}_{12}$, and $\bar{R}_{11}^{\mathsf{C}}$ are nonnegative, and that this rate allocation is valid.

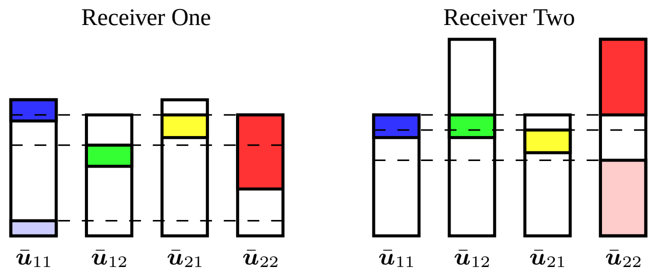

Fig. 12. Allocation of bits in case IV. Here $n_{11} = 18, n_{22} = 26, n_{12} = n_{21} = 16$. The transmitters send private messages at rates $\bar{R}_{11}^{\mathsf{P}} = 2$ and $\bar{R}_{22}^{\mathsf{P}} = 10$. Transmitter one sends a common message to receiver one at rate $\bar{R}_{11}^{\mathsf{C}} = 3$, and transmitter two sends a common message to receiver two at rate $\bar{R}_{22}^{\mathsf{C}} = 10$. The rates over the cross links are $\bar{R}_{12} = 3$ and $\bar{R}_{21} = 3$. In case IV, the interference terms are completely aligned at receiver two, but only partially aligned at receiver one.

The calculation in Appendix A verifies the decoding conditions (19) and (20) in Lemma 7. The sum rate of this allocation is at least

$$\begin{aligned}
&(n_{12} + n_{21} - \tfrac{1}{2}n_{22}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) - 3 \\
&\qquad = \tfrac{1}{2}\big(n_{12} + n_{21} + (n_{12} + n_{21} - n_{22})^{+}\big) + (n_{11} - n_{21}) + (n_{22} - n_{12}) - 3 \\
&\qquad = \bar{D}_2(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) - 3 \\
&\qquad \geq \bar{D}(\boldsymbol{N}) - 3,
\end{aligned} \tag{25}$$

where we have used (24), and where the loss of three bits is due to the floor operation in the definition of $\bar{R}_{21}, \bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}$.

Case V ($\tfrac{3}{2}n_{22} < n_{12} + n_{21} \leq n_{11} + n_{22}$): We set

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11} - n_{21}, \\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22} - n_{12}, \\
\bar{R}_{12} &\triangleq \bar{R}_{11}^{\mathsf{C}} \triangleq \left\lfloor \tfrac{2}{3}n_{21} - \tfrac{1}{3}n_{12} \right\rfloor, \\
\bar{R}_{21} &\triangleq \bar{R}_{22}^{\mathsf{C}} \triangleq \left\lfloor \tfrac{2}{3}n_{12} - \tfrac{1}{3}n_{21} \right\rfloor,
\end{aligned}$$

as shown in Fig. 13. Using $\tfrac{3}{2}n_{22} < n_{12} + n_{21}$, we have

$$2n_{12} - n_{21} \geq 3(n_{22} - n_{21}) \geq 0, \tag{26}$$

implying that $\bar{R}_{21}$ and $\bar{R}_{22}^{\mathsf{C}}$ are nonnegative. Similarly,

$$2n_{21} - n_{12} \geq 3(n_{11} - n_{12}) \geq 0, \tag{27}$$

implying that $\bar{R}_{12}$ and $\bar{R}_{11}^{\mathsf{C}}$ are nonnegative. Hence this rate allocation is valid.

The calculation in Appendix A verifies the decoding conditions (19) and (20) in Lemma 7. The sum rate of this bit allocation is at least

$$\begin{aligned}
&\tfrac{2}{3}(n_{12} + n_{21}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) - 4 \\
&\qquad = \bar{D}_4(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) - 4 \\
&\qquad \geq \bar{D}(\boldsymbol{N}) - 4,
\end{aligned} \tag{28}$$

where the loss of four bits is due to floor operation in the definition of $\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}, \bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}$.

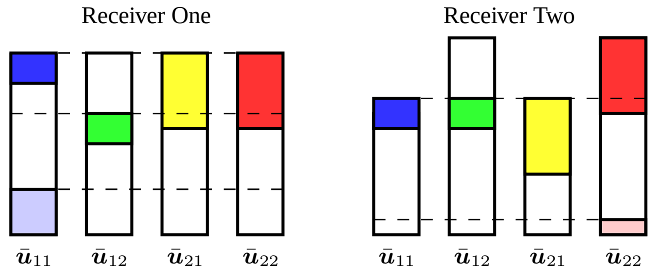

Fig. 13. Allocation of bits in case V. Here $n_{11} = 12, n_{22} = 13, n_{12} = 12, n_{21} = 9$. The private messages to receiver one and two have rates $\bar{R}_{11}^{\mathsf{P}} = 3$ and $\bar{R}_{22}^{\mathsf{P}} = 1$. The remaining messages to receiver one have rate $\bar{R}_{12} = \bar{R}_{11}^{\mathsf{C}} = 2$, and are both entirely aligned at receiver two. The remaining messages to receiver two have rate $\bar{R}_{21} = \bar{R}_{22}^{\mathsf{C}} = 5$, and are both entirely aligned at receiver one.

Combining (21), (22), (23), (25), (28), and accounting for the loss of $2\log(32/\delta)$ in Lemma 7 shows that, assuming $n_{22} \geq n_{11}$,

$$\begin{aligned} \bar{C}(\boldsymbol{N}) &\geq \bar{D}(\boldsymbol{N}) - 4 - 2\log(32/\delta) \\ &= \bar{D}(\boldsymbol{N}) - 2\log(c_1/\delta) \end{aligned}$$

with

$$c_1 \triangleq 128.$$

If $n_{11} \geq n_{22}$, we can simply relabel the two receivers, and the same argument holds. This relabeling of receivers introduces the function $\bar{D}_3(\boldsymbol{N})$ instead of $\bar{D}_2(\boldsymbol{N})$ in the lower bound. Together, this concludes the proof of the lower bound in Theorem 5. ■

B. Upper Bound for the Deterministic X-Channel

The section contains the proof of the upper bound in Theorem 5. We start with a lemma upper bounding various linear combinations of achievable rates for the deterministic X-channel.

Lemma 8. *Any achievable rate tuple* $(\bar{R}_{11}, \bar{R}_{12}, \bar{R}_{21}, \bar{R}_{22})$ *for the (modulated) deterministic X-channel* (17) *satisfies the following inequalities*

$$\begin{aligned}
\bar{R}_{11} + \bar{R}_{12} + \bar{R}_{22} &\leq \max\{n_{11}, n_{12}\} + (n_{22} - n_{12})^+, && \text{(29a)} \\
\bar{R}_{11} + \bar{R}_{21} + \bar{R}_{22} &\leq \max\{n_{21}, n_{22}\} + (n_{11} - n_{21})^+, && \text{(29b)} \\
\bar{R}_{11} + \bar{R}_{21} + \bar{R}_{12} &\leq \max\{n_{11}, n_{12}\} + (n_{21} - n_{11})^+, && \text{(29c)} \\
\bar{R}_{21} + \bar{R}_{12} + \bar{R}_{22} &\leq \max\{n_{21}, n_{22}\} + (n_{12} - n_{22})^+, && \text{(29d)} \\
\bar{R}_{11} + \bar{R}_{21} + \bar{R}_{12} + \bar{R}_{22} &\leq \max\{n_{12}, (n_{11} - n_{21})^+\} + \max\{n_{21}, (n_{22} - n_{12})^+\}, && \text{(29e)} \\
\bar{R}_{11} + \bar{R}_{21} + \bar{R}_{12} + \bar{R}_{22} &\leq \max\{n_{11}, (n_{12} - n_{22})^+\} + \max\{n_{22}, (n_{21} - n_{11})^+\}, && \text{(29f)} \\
2\bar{R}_{11} + \bar{R}_{12} + \bar{R}_{21} + \bar{R}_{22} &\leq \max\{n_{11}, n_{12}\} + \max\{n_{21}, (n_{22} - n_{12})^+\} + (n_{11} - n_{21})^+, && \text{(29g)} \\
\bar{R}_{11} + 2\bar{R}_{12} + \bar{R}_{21} + \bar{R}_{22} &\leq \max\{n_{11}, n_{12}\} + \max\{n_{22}, (n_{21} - n_{11})^+\} + (n_{12} - n_{22})^+, && \text{(29h)} \\
\bar{R}_{11} + \bar{R}_{12} + 2\bar{R}_{21} + \bar{R}_{22} &\leq \max\{n_{22}, n_{21}\} + \max\{n_{11}, (n_{12} - n_{22})^+\} + (n_{21} - n_{11})^+, && \text{(29i)} \\
\bar{R}_{11} + \bar{R}_{12} + \bar{R}_{21} + 2\bar{R}_{22} &\leq \max\{n_{22}, n_{21}\} + \max\{n_{12}, (n_{11} - n_{21})^+\} + (n_{22} - n_{12})^+, && \text{(29j)}
\end{aligned}$$

The proof of Lemma 8 is reported in Appendix B. Inequalities (29a)–(29f) are based on an argument from [10, Lemma 4.2]. Inequalities (29g)–(29j) are novel.

The upper bounds in Lemma 8 can be understood intuitively as multiple-access bounds in a channel where the receivers are forced to decode certain parts of the interference (see Figs. 5 and 6 in Section IV-B). For example, inequality (29a) corresponds to the multiple-access bound

$$\bar{R}_{11} + \bar{R}_{12} + \bar{R}_{22}^{\mathsf{C}} \leq \max\{n_{11}, n_{12}\} \tag{30}$$

at receiver one, combined with the inequality

$$\bar{R}_{22}^{\mathsf{P}} \leq (n_{22} - n_{12})^{+}.$$

Similarly, inequality (29e) corresponds to the multiple-access bound

$$\bar{R}_{11}^{\mathsf{P}} + \bar{R}_{12} + \bar{R}_{22}^{\mathsf{C}} \leq \max\{n_{12}, (n_{11} - n_{21})^{+}\}$$

at receiver one, combined with the multiple-access bound

$$\bar{R}_{22}^{\mathsf{P}} + \bar{R}_{21} + \bar{R}_{11}^{\mathsf{C}} \leq \max\{n_{21}, (n_{22} - n_{12})^{+}\} \tag{31}$$

at receiver two. Finally, inequality (29g) corresponds to the multiple-access bounds (30) and

$$\bar{R}_{11}^{\mathsf{P}} \leq (n_{11} - n_{21})^{+}$$

at receiver one, combined with the multiple-access bound (31) at receiver two. The proof of Lemma 8 makes this intuitive reasoning precise. A detailed discussion of this type of cut-set interpretation can be found in [19].

We proceed with the proof of the upper bound in Theorem 5 for the deterministic X-channel. Under the assumption

$$\min\{n_{11}, n_{22}\} \geq \max\{n_{12}, n_{21}\}, \tag{32}$$

the first four inequalities (29a)–(29d) in Lemma 8 yield the following upper bound on sum-capacity

$$\begin{aligned} \bar{C}(\boldsymbol{N}) &\leq \tfrac{2}{3}(n_{12} + n_{21}) + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\ &= \bar{D}_4(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}). \end{aligned} \tag{33}$$

Again using (32), inequality (29e) in Lemma 8 shows that

$$\begin{aligned} \bar{C}(\boldsymbol{N}) &\leq \max\{n_{12}, n_{11} - n_{21}\} + \max\{n_{21}, n_{22} - n_{12}\} \\ &= (n_{12} + n_{21} - n_{11})^{+} + (n_{12} + n_{21} - n_{22})^{+} + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\ &= \bar{D}_1(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}). \end{aligned} \tag{34}$$

Inequalities (29d) and (29g) in Lemma 8 combined with (32) yield

$$\begin{aligned} \bar{C}(\boldsymbol{N}) &\leq \tfrac{1}{2}\big(n_{11} + n_{22} + \max\{n_{21}, n_{22} - n_{12}\} + (n_{11} - n_{21})\big) \\ &= \tfrac{1}{2}\big(n_{12} + n_{21} + (n_{12} + n_{21} - n_{22})^{+}\big) + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\ &= \bar{D}_2(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}). \end{aligned} \tag{35}$$

Similarly, from (29c) and (29j) in Lemma 8,

$$\begin{aligned} \bar{C}(\boldsymbol{N}) &\leq \tfrac{1}{2}\big(n_{12} + n_{21} + (n_{12} + n_{21} - n_{11})^{+}\big) + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\ &= \bar{D}_3(\boldsymbol{N}) + (n_{11} - n_{21}) + (n_{22} - n_{12}). \end{aligned} \tag{36}$$

The sum capacity is hence at most the minimum of the upper bounds (33)–(36), i.e.,

$$\begin{aligned} \bar{C}(\boldsymbol{N}) &\leq \min\big\{\bar{D}_1(\boldsymbol{N}), \bar{D}_2(\boldsymbol{N}), \bar{D}_3(\boldsymbol{N}), \bar{D}_4(\boldsymbol{N})\big\} + (n_{11} - n_{21}) + (n_{22} - n_{12}) \\ &= \bar{D}(\boldsymbol{N}), \end{aligned}$$

concluding the proof. ■

VI. Proof of Theorem 6 (Gaussian X-Channel)

This section contains the proof of the capacity approximation for the Gaussian X-channel in Theorem 6. Achievability of the lower bound in the theorem is proved in Section VI-A, the upper bound is proved in Section VI-B.

A. Achievability for the Gaussian X-Channel

Here, we prove the lower bound in Theorem 6. For ease of exposition, we assume in most of the analysis that all channels gains h_{mk} are exactly known at the two transmitters and receivers. The changes in the arguments necessary for the mismatched case, in which the transmitters and receivers have access only to a quantized version $\hat{h}_{mk}$ of the channel gain h_{mk}, are reported in Appendix C.

Recall that each transmitter k has access to two messages, w_{1k} and w_{2k}. The transmitter forms the modulated symbol u_{mk} from the message w_{mk}. From these modulated signals, the channel inputs

$$x_1 \triangleq h_{22}u_{11} + h_{12}u_{21},$$
$$x_2 \triangleq h_{21}u_{12} + h_{11}u_{22}$$

are constructed.

We now describe in detail the modulation process from w_{1k} to u_{1k}. Each u_{mk} is of the form

$$u_{mk} \triangleq \sum_{i=3}^{n_{kk}} [u_{mk}]_i 2^{-i}$$

with $[u_{mk}]_i \in \{0, 1\}$. Since $h_{mk}^2 \leq 4$ and $u_{mk}^2 \leq 1/16$, the resulting channel input x_k satisfies the unit average power constraint at the transmitters. Observe that $2^{n_{kk}} u_{mk} \geq 1$; intuitively, this means that over the stronger direct link we only transmit information above the "noise level".

In analogy to the achievable scheme for the deterministic channel, we only use certain portions of the bits $[u_{mk}]_i$ of the binary expansion of u_{mk}; the remaining bits are set to zero. The allocation of information bits depends on the channel strength $\boldsymbol{N}$ and is chosen as in the deterministic case described in Sections IV-B and V-A, and as illustrated in Figs. 8 and 9. In particular, the messages u_{mk} are again decomposed into common and private portions, i.e.,

$$u_{mk} = u_{mk}^{\mathsf{P}} + u_{mk}^{\mathsf{C}}.$$

The choice of bit allocation from the deterministic model ensures that, at receiver one,

$$2^{n_{11}} u_{11} \in \{0, \pm 1, \dots, \pm(2^{n_{11}} - 1)\}, \tag{37a}$$
$$2^{n_{12}} u_{12}^{\mathsf{C}} \in \{0, \pm 1, \dots, \pm(2^{n_{12}} - 1)\}, \tag{37b}$$
$$2^{n_{21}} u_{11}^{\mathsf{C}} \in \{0, \pm 1, \dots, \pm(2^{n_{11}} - 1)\}, \tag{37c}$$
$$2^{n_{12}} u_{22}^{\mathsf{C}} \in \{0, \pm 1, \dots, \pm(2^{n_{12}} - 1)\}, \tag{37d}$$

see Figs. 5 and 6 in Section III-C. The situation is analogous at receiver two. We denote by $\bar{R}_{mk}$ the modulation rate of u_{mk} in bits per symbol in analogy to the deterministic case.

From the bit allocation for the deterministic X-channel, we see that the private portions of the messages u_{12} and u_{21} are always zero, i.e., $u_{12}^{\mathsf{P}} = u_{21}^{\mathsf{P}} = 0$. To satisfy the power constraint (as discussed in the last paragraph), we impose that the two most significant bits of each common message are zero. For reasons that will become clear in the next paragraph, we also impose that the two most significant bit for each private message is zero. This reduces the modulation rate by at most 12 bits per channel use compared to the deterministic case.

The channel output at receiver one is

$$\begin{aligned} y_1 &= 2^{n_{11}} h_{11} x_1 + 2^{n_{12}} h_{12} x_2 + z_1 \\ &= \big(h_{11}h_{22}2^{n_{11}}u_{11} + h_{12}h_{21}2^{n_{12}}u_{12}^{\mathsf{C}}\big) + h_{12}h_{11}\big(2^{n_{11}}u_{21}^{\mathsf{C}} + 2^{n_{12}}u_{22}^{\mathsf{C}}\big) + \big(h_{12}h_{11}2^{n_{12}}u_{22}^{\mathsf{P}} + z_1\big), \end{aligned}$$

and similarly at receiver two. The channel output is grouped into three parts. The first part contains the two desired signals u_{11} and u_{12}^{C}. The second part contains the interference signals u_{21}^{C} and u_{22}^{C}. Note that

these interference terms are received with the same coefficient $h_{12}h_{11}$ and are hence aligned. The third part contains noise z_1 and the private portion u_{22}^{P} of the message u_{22}^{P}. By construction,

$$2^{n_{12}}u_{22}^{\mathsf{P}} \in [0, 1/4)$$

so that

$$|h_{12}h_{11}2^{n_{12}}u_{22}^{\mathsf{P}}| \leq 4 \cdot \tfrac{1}{4} \leq 1. \tag{38}$$

We will treat this part of the interference as noise.

The demodulator at receiver one searches for

$$\begin{aligned}
\hat{s}_{11} &\triangleq 2^{n_{11}}\hat{u}_{11}, \\
\hat{s}_{12} &\triangleq 2^{n_{12}}\hat{u}_{12}^{\mathsf{C}}, \\
\hat{s}_{10} &\triangleq 2^{n_{11}}\hat{u}_{21}^{\mathsf{C}} + 2^{n_{12}}\hat{u}_{22}^{\mathsf{C}}
\end{aligned}$$

that minimizes

$$|y_1 - h_{11}h_{22}\hat{s}_{11} + h_{12}h_{21}\hat{s}_{12} + h_{12}h_{11}\hat{s}_{10}|.$$

Observe that by (37) the search over $\hat{s}_{mk}$ can be restricted to be over integers with magnitude less than $2^{n_{11}}$. The demodulator then declares the minimizing $2^{-n_{11}}\hat{s}_{11}$ and $2^{-n_{12}}\hat{s}_{12}$ as the demodulated symbols $\hat{u}_{11}$ and $\hat{u}_{12}$, respectively, and discards $\hat{s}_{10}$. We point out that the demodulator decodes only the *sum* $\hat{s}_{10}$ of the two interfering symbols, but not the individual interfering symbols themselves. The demodulator at receiver two works in analogy.

We now analyze the probability of error of this demodulator. Let d be the minimum distance between any two possible noiseless received signals resulting from different values of (s_{10}, s_{11}, s_{12}). The probability of error of the demodulator at receiver one is then upper bounded by

$$\begin{aligned}
\mathbb{P}\big(\hat{u}_{1k} \neq u_{1k} \text{ for } k \in \{1,2\}\big) &\leq 2\mathbb{P}(z_1 + |h_{12}h_{11}2^{n_{12}}u_{22}^{\mathsf{P}}| \geq d/2) \\
&\leq 2\mathbb{P}(z_1 \geq d/2 - 1),
\end{aligned} \tag{39}$$

where we have used (38).

We now lower bound the minimum distance d between the received signal generated by the correct (s_{11}, s_{12}, s_{10}) and by any other triple $(s'_{11}, s'_{12}, s'_{10})$. We have

$$\begin{aligned}
d &\triangleq \min_{\substack{(s_{11},s_{12},s_{10}) \\ \neq (s'_{11},s'_{12},s'_{10})}} \big|h_{11}h_{22}(s_{11} - s'_{11}) + h_{12}h_{21}(s_{12} - s'_{12}) + h_{12}h_{11}(s_{10} - s'_{10})\big| \\
&= \min_{\substack{(s_{11},s_{12},s_{10}) \\ \neq (s'_{11},s'_{12},s'_{10})}} \big|g_{11}(s_{11} - s'_{11}) + g_{12}(s_{12} - s'_{12}) + g_{10}(s_{10} - s'_{10})\big|,
\end{aligned}$$

where g_{mk} denotes the products of two h_{mk} as defined in (13). The next lemma provides a sufficient condition for this minimum distance to be large at both receivers.

Lemma 9. *Let* $\delta \in (0,1]$ *and* $\boldsymbol{N} \in \mathbb{Z}_+^{2\times 2}$ *such that* $\min\{n_{11}, n_{22}\} \geq \max\{n_{12}, n_{21}\}$. *Assume* $\bar{R}_{11}^{\mathsf{P}}$, $\bar{R}_{11}^{\mathsf{C}}$, $\bar{R}_{12}$, $\bar{R}_{21}$, $\bar{R}_{22}^{\mathsf{P}}$, $\bar{R}_{22}^{\mathsf{C}} \in \mathbb{Z}_+$ *satisfy*

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{C}} + \max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} &\leq n_{11} - 6 - \log(4368/\delta), \\
\mathbb{1}_{\{\max\{\bar{R}_{12},\bar{R}_{21},\bar{R}_{22}^{\mathsf{C}}\}>0\}} \cdot \big(\max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}}\big) &\leq n_{12} - 6 - \log(4368/\delta), \\
\mathbb{1}_{\{\bar{R}_{12}>0\}} \cdot \big(\bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}}\big) &\leq n_{12} + n_{21} - n_{22} - 6,
\end{aligned}$$

and

$$\begin{aligned}
\bar{R}_{22}^{\mathsf{C}} + \max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}} &\leq n_{22} - 6 - \log(4368/\delta), \\
\mathbb{1}_{\{\max\{\bar{R}_{21},\bar{R}_{12},\bar{R}_{11}^{\mathsf{C}}\}>0\}} \cdot \big(\max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}}\big) &\leq n_{21} - 6 - \log(4368/\delta), \\
\mathbb{1}_{\{\bar{R}_{21}>0\}} \cdot \big(\bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}}\big) &\leq n_{12} + n_{21} - n_{11} - 6.
\end{aligned}$$

Then the bit allocation in Section IV-B applied to the Gaussian X-channel (16) *results in a minimum constellation distance $d \geq 32$ at each receiver for all channel gains $(h_{mk}) \in (1,2]^{2\times 2}$ except for a set $B \subset (1,2]^{2\times 2}$ of Lebesgue measure*

$$\mu(B) \leq \delta.$$

The proof of Lemma 9 is reported in Section VII-B. Observe that, up to the constants, Lemma 9 is exactly of the same form as Lemma 7 in Section V-A for the deterministic X-Channel.

Recall that we have chosen the same allocation of information bits in the binary expansion of u_{mk} as in the deterministic case analyzed in Section V-A. Since the most significant bit of each u_{mk} is zero, the binary expansion of s_{mk} is also of the form analyzed there. Moreover, since the conditions in Lemma 9 used here are the same as the conditions in Lemma 7 used in the deterministic case, we conclude that Lemma 9 can be applied if we further reduce the rates to accommodate the constant $6 + \log(4368/\delta)$ in Lemma 9. This can be achieved for example by reducing the modulation rate by a further $3 + \frac{1}{2}\log(4368/\delta) \leq 9.25 + \frac{1}{2}\log(1/\delta)$ per symbol. Accounting for the 12 bits loss per channel use due to the power constraint, the sum rate of the modulation scheme is then

$$\begin{aligned} \sum_{m,k} \bar{R}_{mk} &= \bar{D}(\boldsymbol{N}) - 12 - 4 \cdot 9.25 - 4 \cdot \tfrac{1}{2}\log(1/\delta) - 4 \\ &= \bar{D}(\boldsymbol{N}) - 2\log(1/\delta) - 53, \end{aligned} \tag{40}$$

with $\bar{D}(\boldsymbol{N})$ as defined in Theorem 5 for the deterministic X-channel, and where the additional loss of 4 bits results from rounding in the bit allocation for the deterministic scheme as discussed in Section V-A.

Lemma 9 yields then that $d \geq 32$ for all h_{mk} except for a set B of measure at most δ. Together with (39), this shows that the probability of demodulation error is upper bounded by

$$\begin{aligned} 2\mathbb{P}(z_1 \geq d/2 - 1) &\leq 2\mathbb{P}(z_1 \geq 15) \\ &\leq \exp(-15^2/2). \end{aligned}$$

As mentioned before, this upper bound on the probability of demodulation error is based on the assumption that both transmitters and receivers have access to h_{mk}. The analysis in Appendix C shows that the only difference under mismatched encoding and decoding, in which the transmitters and receivers use $\max\{n_{mk}\}$-bit quantized channel gains $\hat{h}_{mk}$ instead of h_{mk}, is a decrease in the minimum constellation distance d. In particular, by (89) in Appendix C, the probability of demodulation error with mismatched encoders and decoders is upper bounded by

$$\begin{aligned} 2\mathbb{P}(z_1 \geq d/2 - 7) &\leq 2\mathbb{P}(z_1 \geq 9) \\ &\leq \exp(-9^2/2). \end{aligned}$$

More generally, consider the correct decoding region, and label the incorrect decoding regions of one of the two desired symbols, say s_{11}, in increasing order of distance to the correct one starting from $\ell \geq 1$ (see Fig. 14). By (90) in Appendix C, the probability of decoding to the ℓth such incorrect decoding region with mismatched encoders and decoders is upper bounded by

$$\begin{aligned} \mathbb{P}(z_1 \geq \ell(d-8)/4 - 3) &\leq \mathbb{P}(z_1 \geq 6\ell - 3) \\ &\leq \tfrac{1}{2}\exp(-(6\ell - 3)^2/2). \end{aligned} \tag{41}$$

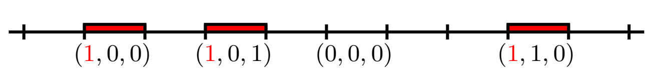

Fig. 14. Decision regions at receiver one. Here, we assume the correct decision is $(s_{11}, s_{12}, s_{10}) = (0,0,0)$. The figure depicts the region resulting in the incorrect decision $\hat{s}_{11} = 1$. This region is a union of intervals, with each interval corresponding to a decision $(1, s'_{12}, s'_{10})$ for some s'_{12}, s'_{10}. See also Fig. 4 in Section II-C.

To achieve a vanishing probability of error for a fixed finite power P, we use an outer code over the modulated channel. Let R_{mk} denote the rate of this outer code from transmitter k to receiver m. The rate R_{1k} achievable from transmitter k to receiver one can be lower bounded by following the first few steps in the proof of Fano's inequality:

$$\begin{aligned} R_{1k} &= I(u;\hat{u}) \\ &= H(u) - H(u \mid \hat{u}) \\ &\geq \bar{R}_{1k} - H(\hat{u} \neq u) - \mathbb{P}(\hat{u} \neq u) H(u \mid \hat{u}, \hat{u} \neq u) \\ &\geq \bar{R}_{1k} - 1 - \mathbb{P}(\hat{u} \neq u) H(u \mid \hat{u}, \hat{u} \neq u), \end{aligned} \tag{42}$$

where we have used u and $\hat{u}$ for u_{1k} and $\hat{u}_{1k}$, respectively.

Remark: The proof of Fano's inequality continues by upper bounding

$$H(u \mid \hat{u}, \hat{u} \neq u) \leq \bar{R}_{1k}.$$

This is the approach taken, for example, in [3], [17]. Since $\bar{R}_{1k}$ itself depends on $\boldsymbol{N}$, this last upper bound is not strong enough to obtain a constant-gap approximation of capacity. Instead, we show next that $H(u \mid \hat{u}, \hat{u} \neq u)$ can be upper bounded by a constant independent of $\boldsymbol{N}$. This argument is one of the key steps in the derivation of the upper bound.

We have

$$\mathbb{P}(\hat{u} \neq u) H(u \mid \hat{u}, \hat{u} \neq u) = \mathbb{P}(\hat{u} \neq u) \sum_{\hat{q}} \mathbb{P}(\hat{u} = \hat{q} \mid \hat{u} \neq u) H(u \mid \hat{u} = \hat{q}, \hat{u} \neq u). \tag{43}$$

Define the probability

$$p_{q,\hat{q}} \triangleq \mathbb{P}(u = q \mid \hat{u} = \hat{q}, \hat{u} \neq u).$$

Applying [21, Theorem 9.7.1],

$$\begin{aligned} H(u \mid \hat{u} = \hat{q}, \hat{u} \neq u) &= -\sum_{q} p_{q,\hat{q}} \log p_{q,\hat{q}} \\ &\leq \frac{1}{2} \log \left((2\pi e) \Big(\sum_{\ell=1}^{2^{\bar{R}_{1k}}} \ell^2 p_{\sigma(\ell),\hat{q}} - \Big(\sum_{\ell=1}^{2^{\bar{R}_{1k}}} \ell p_{\sigma(\ell),\hat{q}} \Big)^2 + \frac{1}{12} \Big) \right) \\ &\leq \frac{1}{2} \log \left((2\pi e) \Big(\sum_{\ell=1}^{2^{\bar{R}_{1k}}} \ell^2 p_{\sigma(\ell),\hat{q}} + \frac{1}{12} \Big) \right) \\ &\leq \frac{1}{2} \log(2\pi e) + \frac{1}{2} \log(e) \sum_{\ell=1}^{2^{\bar{R}_{1k}}} \ell^2 p_{\sigma(\ell),\hat{q}} \end{aligned} \tag{44}$$

for any invertible map σ into $\{1, \dots, 2^{\bar{R}_{1k}}\}$.

Now,

$$\begin{aligned} p_{q,\hat{q}} &= \frac{\mathbb{P}(u = q, \hat{u} = \hat{q}, \hat{u} \neq u)}{\mathbb{P}(\hat{u} = \hat{q}, \hat{u} \neq u)} \\ &= \frac{\mathbb{P}(u = q) \mathbb{P}(\hat{u} = \hat{q} \mid u = q) \mathbb{P}(\hat{u} \neq u \mid u = q, \hat{u} = \hat{q})}{\mathbb{P}(\hat{u} \neq u) \mathbb{P}(\hat{u} = \hat{q} \mid \hat{u} \neq u)} \\ &= \frac{\mathbb{P}(\hat{u} = \hat{q} \mid u = q) \mathbb{1}_{\hat{q} \neq q}}{2^{\bar{R}_{1k}} \mathbb{P}(\hat{u} \neq u) \mathbb{P}(\hat{u} = \hat{q} \mid \hat{u} \neq u)}. \end{aligned}$$

Combining this with (44) yields

$$H(u \mid \hat{u}=\hat{q}, \hat{u}\neq u) \leq \frac{1}{2}\log(2\pi e) + \frac{1}{2}\log(e)\frac{\sum_{\ell=1}^{2^{\bar{R}_{1k}}} \ell^2 \mathbb{1}_{\sigma(\ell)\neq\hat{q}} \mathbb{P}(\hat{u}=\hat{q} \mid u=\sigma(\ell))}{2^{\bar{R}_{1k}}\mathbb{P}(\hat{u}\neq u)\mathbb{P}(\hat{u}=\hat{q}\mid \hat{u}\neq u)}.$$

Choose the map σ such that $\sigma(1)=\hat{q}$. Then this last expression can be simplified to

$$H(u \mid \hat{u}=\hat{q}, \hat{u}\neq u) \leq \frac{1}{2}\log(2\pi e) + \frac{1}{2}\log(e)\frac{\sum_{\ell=2}^{2^{\bar{R}_{1k}}} \ell^2 \mathbb{P}(\hat{u}=\hat{q} \mid u=\sigma(\ell))}{2^{\bar{R}_{1k}}\mathbb{P}(\hat{u}\neq u)\mathbb{P}(\hat{u}=\hat{q}\mid \hat{u}\neq u)}. \tag{45}$$

Consider the decision regions at the decoder that result in the incorrect decision $\hat{u}=\hat{u}_{1k}=\hat{q}$. This region is a union of intervals with each interval corresponding to a different triple $(s'_{11}, s'_{12}, s'_{10})$, see Fig. 14. The probability of decoding $\hat{u}_{1k}=\hat{q}$ can hence be upper bounded by the probability that the magnitude of the noise is larger than the distance between the noiseless channel output for the transmitted triple (s_{11}, s_{12}, s_{12}) and the closest of the intervals of the decision region resulting in $\hat{u}_{1k}=\hat{q}$. This probability is upper bounded by (41), showing that we can choose σ such that, for $\ell \geq 2$,

$$\begin{aligned}\mathbb{P}(\hat{u}=\hat{q} \mid u=\sigma(\ell)) &\leq 2\mathbb{P}(z_1 \geq 6(\ell-1)-3) \\ &\leq \exp(-(6\ell-9)^2/2),\end{aligned}$$

so that

$$\sum_{\ell=2}^{2^{\bar{R}_{1k}}} \ell^2 \mathbb{P}(\hat{u}=\hat{q} \mid u=\sigma(\ell)) \leq \sum_{\ell=2}^{\infty} \ell^2 \exp(-(6\ell-9)^2/2) \leq 1.$$

Together with (43) and (45), this shows that

$$\begin{aligned}&\mathbb{P}(\hat{u}\neq u)H(u \mid \hat{u}, \hat{u}\neq u) \\ &\quad\leq \sum_{\hat{q}} \mathbb{P}(\hat{u}\neq u)\mathbb{P}(\hat{u}=\hat{q}\mid \hat{u}\neq u)\bigg(\frac{1}{2}\log(2\pi e) + \frac{1}{2}\log(e)\frac{1}{2^{\bar{R}_{1k}}\mathbb{P}(\hat{u}\neq u)\mathbb{P}(\hat{u}=\hat{q}\mid \hat{u}\neq u)}\bigg) \\ &\quad\leq \frac{1}{2}\log(2\pi e^2).\end{aligned} \tag{46}$$

Substituting (46) into (42) shows that the rate between transmitter k and receiver one is lower bounded by

$$R_{1k} \geq \bar{R}_{1k} - 1 - \tfrac{1}{2}\log(2\pi e^2) \tag{47}$$

for all $(h_{mk}) \in (1,2]^{2\times 2}$ except for a set B of measure at most δ.

The same argument holds for the rates R_{2k} at the second receiver with the same outage set B. This shows that, for channel gains not in B, we can achieve a sum rate of at least

$$\sum_{m,k} R_{mk} \geq \sum_{m,k} \bar{R}_{mk} - 4 - 2\log(2\pi e^2),$$

and hence

$$C(\boldsymbol{N}) \geq \sum_{m,k} \bar{R}_{mk} - 4 - 2\log(2\pi e^2).$$

Using (40), this shows that, except for a set B of measure δ,

$$\begin{aligned}C(\boldsymbol{N}) &\geq \bar{D}(\boldsymbol{N}) - 2\log(1/\delta) - 57 - 2\log(2\pi e^2) \\ &= D(\boldsymbol{N}) - 2\log(c_2/\delta),\end{aligned}$$

since $\bar{D}(\boldsymbol{N}) = D(\boldsymbol{N})$ by definition, and with

$$c_2 \triangleq 2\pi e^2 \cdot 2^{28.5}.$$

This concludes the proof of the lower bound in Theorem 6. ■

B. Upper Bound for the Gaussian X-Channel

This section proves the upper bound in Theorem 6. We start with a lemma upper bounding various linear combinations of achievable rates for the Gaussian X-channel.

Lemma 10. *Any achievable rate tuple* $(R_{11}, R_{12}, R_{21}, R_{22})$ *for the Gaussian X-channel* (16) *satisfies the following inequalities*

$$R_{11}+R_{12}+R_{22} \leq \tfrac{1}{2}\log(1+2^{2n_{11}}h_{11}^2+2^{2n_{12}}h_{12}^2)+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{22}}h_{22}^2}{1+2^{2n_{12}}h_{12}^2}\right), \tag{48a}$$

$$R_{11}+R_{21}+R_{22} \leq \tfrac{1}{2}\log(1+2^{2n_{22}}h_{22}^2+2^{2n_{21}}h_{21}^2)+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{11}}h_{11}^2}{1+2^{2n_{21}}h_{21}^2}\right), \tag{48b}$$

$$R_{11}+R_{21}+R_{12} \leq \tfrac{1}{2}\log(1+2^{2n_{11}}h_{11}^2+2^{2n_{12}}h_{12}^2)+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{21}}h_{21}^2}{1+2^{2n_{11}}h_{11}^2}\right), \tag{48c}$$

$$R_{21}+R_{12}+R_{22} \leq \tfrac{1}{2}\log(1+2^{2n_{22}}h_{22}^2+2^{2n_{21}}h_{21}^2)+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{12}}h_{12}^2}{1+2^{2n_{22}}h_{22}^2}\right), \tag{48d}$$

$$R_{11}+R_{21}+R_{12}+R_{22} \leq \tfrac{1}{2}\log\left(1+2^{2n_{12}}h_{12}^2+\frac{2^{2n_{11}}h_{11}^2}{1+2^{2n_{21}}h_{21}^2}\right)+\tfrac{1}{2}\log\left(1+2^{2n_{21}}h_{21}^2+\frac{2^{2n_{22}}h_{22}^2}{1+2^{2n_{12}}h_{12}^2}\right), \tag{48e}$$

$$R_{11}+R_{21}+R_{12}+R_{22} \leq \tfrac{1}{2}\log\left(1+2^{2n_{11}}h_{11}^2+\frac{2^{2n_{12}}h_{12}^2}{1+2^{2n_{22}}h_{22}^2}\right)+\tfrac{1}{2}\log\left(1+2^{2n_{22}}h_{22}^2+\frac{2^{2n_{21}}h_{21}^2}{1+2^{2n_{11}}h_{11}^2}\right), \tag{48f}$$

$$\begin{aligned}2R_{11}+R_{12}+R_{21}+R_{22} \leq{}& \tfrac{1}{2}\log(1+2^{2n_{11}}h_{11}^2+2^{2n_{12}}h_{12}^2)+\tfrac{1}{2}\log\left(1+2^{2n_{21}}h_{21}^2+\frac{2^{2n_{22}}h_{22}^2}{1+2^{2n_{12}}h_{12}^2}\right)\\&+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{11}}h_{11}^2}{1+2^{2n_{21}}h_{21}^2}\right),\end{aligned} \tag{48g}$$

$$\begin{aligned}R_{11}+2R_{12}+R_{21}+R_{22} \leq{}& \tfrac{1}{2}\log(1+2^{2n_{12}}h_{12}^2+2^{2n_{11}}h_{11}^2)+\tfrac{1}{2}\log\left(1+2^{2n_{22}}h_{22}^2+\frac{2^{2n_{21}}h_{21}^2}{1+2^{2n_{11}}h_{11}^2}\right)\\&+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{12}}h_{12}^2}{1+2^{2n_{22}}h_{22}^2}\right),\end{aligned} \tag{48h}$$

$$\begin{aligned}R_{11}+R_{12}+2R_{21}+R_{22} \leq{}& \tfrac{1}{2}\log(1+2^{2n_{21}}h_{21}^2+2^{2n_{22}}h_{22}^2)+\tfrac{1}{2}\log\left(1+2^{2n_{11}}h_{11}^2+\frac{2^{2n_{12}}h_{12}^2}{1+2^{2n_{22}}h_{22}^2}\right)\\&+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{21}}h_{21}^2}{1+2^{2n_{11}}h_{11}^2}\right),\end{aligned} \tag{48i}$$

$$\begin{aligned}R_{11}+R_{12}+R_{21}+2R_{22} \leq{}& \tfrac{1}{2}\log(1+2^{2n_{22}}h_{22}^2+2^{2n_{21}}h_{21}^2)+\tfrac{1}{2}\log\left(1+2^{2n_{12}}h_{12}^2+\frac{2^{2n_{11}}h_{11}^2}{1+2^{2n_{21}}h_{21}^2}\right)\\&+\tfrac{1}{2}\log\left(1+\frac{2^{2n_{22}}h_{22}^2}{1+2^{2n_{12}}h_{12}^2}\right).\end{aligned} \tag{48j}$$

The proof of Lemma 10 is reported in Appendix D. Inequalities (48a)–(48f) are from [10, Lemma 5.2, Theorem 5.3]. Inequalities (48g)–(48j) are novel.

We proceed with the proof of the upper bound in Theorem 6 for the Gaussian X-channel. Note that, for $n_{mk} \in \mathbb{Z}_+$ and $h_{mk} \in (1,2]$,

$$\begin{aligned}\tfrac{1}{2}\log\left(1+2^{2n_{11}}h_{11}^2+2^{2n_{12}}h_{12}^2\right) &\leq \tfrac{1}{2}\log\left(1+4\cdot 2^{2n_{11}}+4\cdot 2^{2n_{12}}\right)\\&\leq \tfrac{1}{2}\log\left(9\max\{1,2^{2n_{11}},2^{2n_{12}}\}\right)\\&= \tfrac{1}{2}\log(9)+\max\{n_{11},n_{12}\}\end{aligned}$$

and

$$\tfrac{1}{2}\log\Big(1+\frac{2^{2n_{22}}h_{22}^2}{1+2^{2n_{12}}h_{12}^2}\Big) \leq \tfrac{1}{2}\log\big(1+2^{2n_{22}-2n_{12}}h_{22}^2\big)$$
$$\leq \tfrac{1}{2}\log\big(5\max\{1,2^{2n_{22}-2n_{12}}\}\big)$$
$$= \tfrac{1}{2}\log(5)+(n_{22}-n_{12})^+.$$

In a similar manner, we can upper bound the right-hand sides of all terms in Lemma 10 by quantities depending only on $\boldsymbol{N}$.

Under the assumption

$$\min\{n_{11},n_{22}\} \geq \max\{n_{12},n_{21}\}, \tag{49}$$

the first four inequalities (48a)–(48d) in Lemma 10 yield the following upper bound on the sum-capacity

$$\begin{aligned} C(N) &\leq \tfrac{2}{3}(n_{12}+n_{21})+(n_{11}-n_{21})+(n_{22}-n_{12})+\tfrac{2}{3}\log(45) \\ &\leq D_4(\boldsymbol{N})+(n_{11}-n_{21})+(n_{22}-n_{12})+4. \end{aligned} \tag{50}$$

Again using (49), inequality (48e) in Lemma 10 shows that

$$\begin{aligned} C(\boldsymbol{N}) &\leq \max\{n_{12},n_{11}-n_{21}\}+\max\{n_{21},n_{22}-n_{12}\}+\log(9) \\ &= (n_{12}+n_{21}-n_{11})^+ + (n_{12}+n_{21}-n_{22})^+ + (n_{11}-n_{21})+(n_{22}-n_{12})+\log(9) \\ &\leq D_1(\boldsymbol{N})+(n_{11}-n_{21})+(n_{22}-n_{12})+4. \end{aligned} \tag{51}$$

Inequality (48d) and (48g) in Lemma 10 combined with (49) yield

$$\begin{aligned} C(\boldsymbol{N}) &\leq \tfrac{1}{2}\big(n_{22}+n_{11}+\max\{n_{21},n_{22}-n_{12}\}+(n_{11}-n_{21})\big)+\tfrac{1}{4}\log(5^2\cdot 9^3) \\ &= \tfrac{1}{2}\big(n_{12}+n_{21}+(n_{12}+n_{21}-n_{22})^+\big)+(n_{11}-n_{21})+(n_{22}-n_{12})+\tfrac{1}{4}\log(5^2\cdot 9^3) \\ &\leq D_2(\boldsymbol{N})+(n_{11}-n_{21})+(n_{22}-n_{12})+4. \end{aligned} \tag{52}$$

Similarly, from (48c) and (48j) in Lemma 10,

$$\begin{aligned} C(\boldsymbol{N}) &\leq \tfrac{1}{2}\big(n_{12}+n_{21}+(n_{12}+n_{21}-n_{11})^+\big)+(n_{11}-n_{21})+(n_{22}-n_{12})+\tfrac{1}{4}\log(5^2\cdot 9^3) \\ &\leq D_3(\boldsymbol{N})+(n_{11}-n_{21})+(n_{22}-n_{12})+4. \end{aligned} \tag{53}$$

The sum capacity is hence at most the minimum of the four upper bounds (50)–(53), i.e.,

$$\begin{aligned} C(\boldsymbol{N}) &\leq \min\big\{D_1(\boldsymbol{N}),D_2(\boldsymbol{N}),D_3(\boldsymbol{N}),D_4(\boldsymbol{N})\big\}+(n_{11}-n_{21})+(n_{22}-n_{12})+4 \\ &= D(\boldsymbol{N})+4, \end{aligned}$$

concluding the proof. ∎

VII. Mathematical Foundations for Receiver Analysis

This section lays the mathematical groundwork for the analysis of the decoders used in Sections V-A and VI-A. For the deterministic channel model, decoding is successful if the various message subspaces are linearly independent. Conditions for this linear independence to hold are presented in Section VII-A. For the Gaussian case, decoding is successful if the minimum distance between the different messages as seen at the receiver is large. As we will see, this problem can be reformulated as a number-theoretic problem. Conditions for successful decoding in the Gaussian case are presented in Section VII-B.

A. Decoding Conditions for the Deterministic Channel

We start by analyzing a "generic" receiver (i.e., the bit allocation seen at either receiver one or two). To this end, we assume there are two desired vectors $\bar{\boldsymbol{u}}_1$ and $\bar{\boldsymbol{u}}_2$ and one interference vector $\bar{\boldsymbol{u}}_0$. The interference vector consists of two signal vectors that are aligned, and can therefore be treated as a single vector. These three vectors are multiplied by the lower triangular channel matrices $\bar{\boldsymbol{G}}_1, \bar{\boldsymbol{G}}_2$, and $\bar{\boldsymbol{G}}_0$ created via the binary expansion of the channel gains g_1, g_2, g_0 as before. We assume that certain components of the vectors $\bar{\boldsymbol{u}}_k$ are set to zero. Specifically, we consider vectors in the set

$$\begin{aligned}\bar{\mathcal{U}} \triangleq \big\{\bar{\boldsymbol{u}}_k \in \{0,1\}^{n_1}\ \forall k\{0,1,2,3\} :\\ &\bar{u}_{0i} = 0 \text{ for } i \in \{1,\dots,n_1-n_0\} \cup \{n_1-n_0+\bar{R}_0+1,\dots,n_1\}\\ &\bar{u}_{1i} = 0 \text{ for } i \in \{\bar{R}_1+1,\dots,n_1\},\\ &\bar{u}_{2i} = 0 \text{ for } i \in \{1,\dots,n_1-n_2\} \cup \{n_1-n_2+\bar{R}_2+1,\dots,n_1\},\\ &\bar{u}_{3i} = 0 \text{ for } i \in \{1,\dots,n_1-\bar{R}_3\}\big\}.\end{aligned}$$

Here, $\bar{\boldsymbol{u}}_1$ and $\bar{\boldsymbol{u}}_3$ are to be interpreted as the common and private messages from transmitter one, respectively. This situation is illustrated in Fig. 15. The next lemma states that the subspaces spanned by

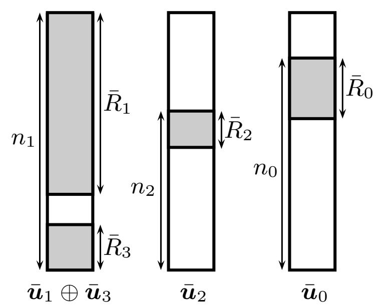

Fig. 15. A generic receiver as analyzed in Lemma 11. White regions correspond to zero bits; shaded regions carry information. Bits are labeled from 1 to n, starting from the top.

the corresponding columns of $\bar{\boldsymbol{G}}_k$ are linearly independent for most channel gains (g_0, g_1, g_2).

Lemma 11. *Let $n_0, n_1, n_2 \in \mathbb{Z}_+$ such that $n_1 \geq n_0 \geq n_2$, and let $\bar{R}_0, \bar{R}_1, \bar{R}_2, \bar{R}_3 \in \mathbb{Z}_+$. Define the event*

$$B(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2, \bar{\boldsymbol{u}}_3) \triangleq \big\{(g_0, g_1, g_2) \in (1,2]^3 : \bar{\boldsymbol{G}}_0\bar{\boldsymbol{u}}_0 \oplus \bar{\boldsymbol{G}}_1(\bar{\boldsymbol{u}}_1 \oplus \bar{\boldsymbol{u}}_3) \oplus \bar{\boldsymbol{G}}_2\bar{\boldsymbol{u}}_2 = 0\big\},$$

and set

$$B \triangleq \bigcup_{(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2, \bar{\boldsymbol{u}}_3) \in \bar{\mathcal{U}} \setminus \{(\boldsymbol{0},\boldsymbol{0},\boldsymbol{0},\boldsymbol{0})\}} B(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2, \bar{\boldsymbol{u}}_3).$$

If there exists $\delta \in (0,1]$ satisfying

$$\begin{aligned}\bar{R}_1 + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_1 - \log(16/\delta),\\ \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_0 - \log(16/\delta),\\ \bar{R}_2 + \bar{R}_3 &\leq n_2,\end{aligned}$$

then

$$\mu(B) \leq \delta.$$

Observe that B is the set of channel gains g_0, g_1, g_2 such that the corresponding subspaces spanned by the selected columns of $\bar{\boldsymbol{G}}_0, \bar{\boldsymbol{G}}_1, \bar{\boldsymbol{G}}_2$ are linearly dependent. Thus the lemma states that if the rates $\bar{R}_k$ satisfy certain conditions, then the subspaces under consideration are linearly independent with high probability.

The condition on the rates in Lemma 11 can be interpreted as follows. Since the matrices $\bar{\boldsymbol{G}}_k$ are lower triangular, the subspaces spanned by the last n columns of $\bar{\boldsymbol{G}}_k$ are the same for all $k \in \{0,1,2\}$. Thus, a *necessary* condition for the linear independence of the three subspaces is that the total number of possible nonzero components of $\bar{u}_{ki}$ with $i \geq n_1 - n + 1$ and $k \in \{0,1,2\}$ is at most n. By the structure of the set $\bar{\mathcal{U}}$, this condition can be verified by considering only three values of n, namely $n \in \{n_0, n_1, n_2\}$. Thus, a necessary condition for the linear independence of the subspaces is

$$\begin{aligned} \bar{R}_1 + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_1, \\ (\bar{R}_1 - (n_1 - n_0))^+ + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_0, \\ (\bar{R}_1 - (n_1 - n_2))^+ + (\bar{R}_0 - (n_0 - n_2))^+ + \bar{R}_2 + \bar{R}_3 &\leq n_2. \end{aligned}$$

After some algebra these three conditions can be rewritten equivalently as

$$\begin{aligned} \bar{R}_1 + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_1, \\ \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_0, \\ \bar{R}_2 + \bar{R}_3 &\leq n_2. \end{aligned}$$

Lemma 11 shows that, up to a constant $\log(16/\delta)$ and for most channel gains (g_0, g_1, g_2), these *necessary* conditions are also *sufficient* for the linear independence of the subspaces.

Before we provide the proof of Lemma 11, we show how it can be used to prove Lemma 7 in Section V-A. We will apply Lemma 11 with δ replaced by $\delta/2$ to guarantee that the outage event at each receiver has measure at most $\delta/2$.

We start by reformulating the conditions in Lemma 11 for each receiver. Consider first receiver one in Lemma 7. The corresponding message rates in Lemma 11 are given by

$$\begin{aligned} \bar{R}_0 &\triangleq \max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\}, \\ \bar{R}_1 &\triangleq \bar{R}_{11}^{\mathsf{C}}, \\ \bar{R}_2 &\triangleq \bar{R}_{12}, \\ \bar{R}_3 &\triangleq \bar{R}_{11}^{\mathsf{P}}. \end{aligned}$$

The choice of the bit levels n_k in Lemma 11 depends on the values of $\bar{R}_0$ and $\bar{R}_2$. If $\bar{R}_0, \bar{R}_2 > 0$, we need to set

$$\begin{aligned} n_0 &\triangleq n_{12}, \\ n_1 &\triangleq n_{11}, \\ n_2 &\triangleq n_{12} + n_{21} - n_{22}, \end{aligned}$$

see Fig. 8 in Section IV-B. The conditions in Lemma 11 are then that

$$\bar{R}_{11}^{\mathsf{C}} + \max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11} - \log(32/\delta), \tag{54a}$$

$$\max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{12} - \log(32/\delta), \tag{54b}$$

$$\bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{12} + n_{21} - n_{22}. \tag{54c}$$

If $\bar{R}_0 > 0$, $\bar{R}_2 = 0$, then the second column in Fig. 15 is empty, and hence the third condition in Lemma 11 need not be verified. Formally, note that in this case the value of n_2 is irrelevant to the decoding process. We may hence assume without loss of generality that n_2 is equal to n_0 (thus still satisfying $n_0 \geq n_2$). Thus only conditions (54a) and (54b) need to be checked. If $\bar{R}_0 = \bar{R}_2 = 0$, the values of n_0 and n_2 are irrelevant to the decoding process, and we can assume them to be equal to n_1 (again still satisfying $n_1 \geq n_0 \geq n_2$). Thus only condition (54a) needs to be checked.

Consider now receiver two in Lemma 7. By symmetry, Lemma 11, results in the decoding conditions

$$\bar{R}_{22}^{\mathsf{C}} + \max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}} \leq n_{22} - \log(32/\delta), \tag{55a}$$

$$\mathbb{1}_{\{\max\{\bar{R}_{21}, \bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\}>0\}} \cdot \big(\max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}}\big) \leq n_{21} - \log(32/\delta), \tag{55b}$$

$$\mathbb{1}_{\{\bar{R}_{21}>0\}} \cdot \big(\bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}}\big) \leq n_{12} + n_{21} - n_{11}. \tag{55c}$$

Denote by $B_1 \subseteq \mathbb{R}^3$ the collection of triples (g_{10}, g_{11}, g_{12}) such that decoding fails at receiver one. Similarly, define $B_2 \subseteq \mathbb{R}^3$ with respect to receiver two. Finally, let $B \subset \mathbb{R}^6$ be the union of B_1 and B_2. If the two sets of decoding conditions (54) and (55) are satisfied, then Lemma 11 shows that

$$\mu_3(B_m) \leq \delta/2$$

for $m \in \{1, 2\}$, where here and in the following we use the notation μ_d to emphasize the Lebesgue measure is computed in $\mathbb{R}^d$. Then

$$\begin{aligned} \mu_6(B) &\leq \mu_6(B_1) + \mu_6(B_2) \\ &= \mu_3(B_1) \cdot \mu_3((1,2]^3) + \mu_3(B_2) \cdot \mu_3((1,2]^3) \\ &\leq \delta, \end{aligned}$$

i.e., the collection of channel gains $(g_{mk}) \in \mathbb{R}^{2\times 3}$ for which decoding fails is small. This concludes the proof of Lemma 7.

It remains to prove Lemma 11.

Proof of Lemma 11: Throughout this proof, we will use the notation $n(\bar{\boldsymbol{u}})$ to denote the smallest index i such that $\bar{u}_i = 1$, with the convention that $n(\mathbf{0}) = +\infty$. Observe that, by the assumptions on $\bar{R}_k$,

$$\begin{aligned} \bar{R}_1 + \bar{R}_3 &\leq n_1, \\ \bar{R}_0 + \bar{R}_3 &\leq n_0, \\ \bar{R}_2 + \bar{R}_3 &\leq n_2, \end{aligned}$$

which implies that

$$\max\big\{\bar{R}_1, n_1 - n_0 + \bar{R}_0, n_1 - n_2 + \bar{R}_2\big\} \leq n_1 - \bar{R}_3. \tag{56}$$

By the definition of $\bar{\mathcal{U}}$, this guarantees that, if $\bar{\boldsymbol{u}}_k \neq 0$, then $n(\bar{\boldsymbol{u}}_k) \leq n_1 - \bar{R}_3$ for $k \in \{0, 1, 2\}$. Moreover, since $\bar{\boldsymbol{G}}_k$ is lower triangular with unit diagonal, we have $n(\bar{\boldsymbol{G}}_k \bar{\boldsymbol{u}}_k) = n(\bar{\boldsymbol{u}}_k)$. Together this implies that

$$\bar{\boldsymbol{G}}_0 \bar{\boldsymbol{u}}_0 \oplus \bar{\boldsymbol{G}}_1(\bar{\boldsymbol{u}}_1 \oplus \bar{\boldsymbol{u}}_3) \oplus \bar{\boldsymbol{G}}_2 \bar{\boldsymbol{u}}_2 = \mathbf{0} \tag{57}$$

only if

$$n\big(\bar{\boldsymbol{G}}_0 \bar{\boldsymbol{u}}_0 \oplus \bar{\boldsymbol{G}}_1 \bar{\boldsymbol{u}}_1 \oplus \bar{\boldsymbol{G}}_2 \bar{\boldsymbol{u}}_2\big) > n_1 - \bar{R}_3.$$

By the same argument, assuming $\bar{\boldsymbol{u}}_3 \neq 0$, (57) can hold only if $(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2) \neq (\mathbf{0}, \mathbf{0}, \mathbf{0})$. Defining the sets

$$\bar{\mathcal{U}}' \triangleq \big\{(\bar{\boldsymbol{u}}_k \in \{0,1\}^{n_1} \ \forall k\{0,1,2\} : (\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2, \mathbf{0}) \in \bar{\mathcal{U}}\big\}$$

and

$$B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2) \triangleq \big\{(g_0, g_1, g_2) \in (1,2]^3 : n\big(\bar{\boldsymbol{G}}_0 \bar{\boldsymbol{u}}_0 \oplus \bar{\boldsymbol{G}}_1 \bar{\boldsymbol{u}}_1 \oplus \bar{\boldsymbol{G}}_2 \bar{\boldsymbol{u}}_2\big) > n_1 - \bar{R}_3\big\},$$

we hence have

$$B \subseteq B' \triangleq \bigcup_{(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2) \in \bar{\mathcal{U}}' \setminus (\mathbf{0},\mathbf{0},\mathbf{0})} B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2).$$

We can then upper bound $\mu(B)$ using the union bound

$$\mu(B) \leq \mu(B') \leq \sum_{(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2) \in \bar{\mathcal{U}}' \setminus (\mathbf{0},\mathbf{0},\mathbf{0})} \mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big). \tag{58}$$

We continue by analyzing each term in the summation on the right-hand side separately.

Since we are integrating with respect to Lebesgue measure over $(g_0, g_1, g_2) \in (1,2]^3$, we can equivalently assume that g_0, g_1, g_2 are independent and uniformly distributed over $(1,2]$. The bits in the binary expansion $([g_k]_i)_{i=-\infty}^{\infty}$ of these numbers are then binary random variables with the following properties. $[g_k]_i = 0$ for $i \leq -1$, $[g_k]_0 = 1$, and $([g_k]_i)_{i=\infty}^{1}$ are i.i.d. $\mathrm{Bernoulli}(1/2)$ (see, e.g., [22, Exercise 1.4.20]). The matrix $\bar{\boldsymbol{G}}_k$ is then constructed from these binary random variables. Note that this implies that the three matrices $\bar{\boldsymbol{G}}_0, \bar{\boldsymbol{G}}_1, \bar{\boldsymbol{G}}_2$ are independent and identically distributed. Recall that each $\bar{\boldsymbol{G}}_k$ is lower triangular Toeplitz with unit diagonal.

Fix a binary vector $\bar{\boldsymbol{u}}$ and consider the product

$$\boldsymbol{b} \triangleq \bar{\boldsymbol{G}}\bar{\boldsymbol{u}}$$

for some $\bar{\boldsymbol{G}} = \bar{\boldsymbol{G}}_k$, $\bar{\boldsymbol{u}} = \bar{\boldsymbol{u}}_k$, and with addition again over $\mathbb{Z}_2$. We now describe the distribution of $\boldsymbol{b}$. Since $\bar{\boldsymbol{G}}$ is lower triangular with unit diagonal, $b_i = 0$ whenever $1 \leq i < n(\bar{\boldsymbol{u}})$, and $b_{n(\bar{\boldsymbol{u}})} = 1$. In particular, $n(\boldsymbol{b}) = n(\boldsymbol{u})$. Moreover, the components b_i for $n(\bar{\boldsymbol{u}}) < i \leq n_1$ are i.i.d. $\mathrm{Bernoulli}(1/2)$.

Assume first that

$$n(\bar{\boldsymbol{u}}_0) \leq n(\bar{\boldsymbol{u}}_1) \leq n(\bar{\boldsymbol{u}}_2) < \infty. \tag{59}$$

The summand in (58) can be written as

$$\mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) = \sum_{\boldsymbol{b}_3 : n(\boldsymbol{b}_3) \geq n_1 - \bar{R}_3 + 1} \sum_{\boldsymbol{b}_1} \sum_{\boldsymbol{b}_2} \mathbb{P}(\bar{\boldsymbol{G}}_0\bar{\boldsymbol{u}}_0 = \boldsymbol{b}_1 \oplus \boldsymbol{b}_2 \oplus \boldsymbol{b}_3)\mathbb{P}(\bar{\boldsymbol{G}}_1\bar{\boldsymbol{u}}_1 = \boldsymbol{b}_1)\mathbb{P}(\bar{\boldsymbol{G}}_2\bar{\boldsymbol{u}}_2 = \boldsymbol{b}_2). \tag{60}$$

Each term in the summation is nonzero only if

$$\begin{aligned} n(\boldsymbol{b}_1 \oplus \boldsymbol{b}_2 \oplus \boldsymbol{b}_3) &= n(\bar{\boldsymbol{u}}_0), \\ n(\boldsymbol{b}_1) &= n(\bar{\boldsymbol{u}}_1), \\ n(\boldsymbol{b}_2) &= n(\bar{\boldsymbol{u}}_2). \end{aligned}$$

Since $n_1 - \bar{R}_3 + 1$ is strictly larger than $n(\bar{\boldsymbol{u}}_k)$ for all $k \in \{0,1,2\}$ by (56) and using (59), the equality $n(\boldsymbol{b}_1 \oplus \boldsymbol{b}_2 \oplus \boldsymbol{b}_3) = n(\bar{\boldsymbol{u}}_0)$ can hold only if $n(\bar{\boldsymbol{u}}_0) = n(\bar{\boldsymbol{u}}_1)$. If the conditions on the $\bar{\boldsymbol{u}}_k$ and $\boldsymbol{b}_k$ are satisfied, then

$$\mathbb{P}(\bar{\boldsymbol{G}}_0\bar{\boldsymbol{u}}_0 = \boldsymbol{b}_1 \oplus \boldsymbol{b}_2 \oplus \boldsymbol{b}_3)\mathbb{P}(\bar{\boldsymbol{G}}_1\bar{\boldsymbol{u}}_1 = \boldsymbol{b}_1)\mathbb{P}(\bar{\boldsymbol{G}}_2\bar{\boldsymbol{u}}_2 = \boldsymbol{b}_2) = 2^{-(n_1 - n(\bar{\boldsymbol{u}}_0)) - (n_1 - n(\bar{\boldsymbol{u}}_1)) - (n_1 - n(\bar{\boldsymbol{u}}_2))}.$$

Substituting this into (60) shows that

$$\mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) \leq 2^{\bar{R}_3 - (n_1 - n(\bar{\boldsymbol{u}}_0))}$$

whenever $n(\bar{\boldsymbol{u}}_0) = n(\bar{\boldsymbol{u}}_1)$, and

$$\mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) = 0$$

otherwise.

Assume more generally that $(\bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2, \bar{\boldsymbol{u}}_3) \neq (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$. Then a similar argument shows that

$$\mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) \leq 2^{\bar{R}_3 - n_1 + \min_k n(\bar{\boldsymbol{u}}_k)} \tag{61}$$

whenever there are two distinct indices k, k' achieving the minimum $\min_k n(\boldsymbol{u}_k)$, and

$$\mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) = 0 \tag{62}$$

otherwise. In particular, the set $B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)$ has measure zero whenever at least two of the $\bar{\boldsymbol{u}}_k$ are equal to zero.

For $n^- \geq n^+$, let

$$\bar{\mathcal{U}}'(n^-, n^+) \triangleq \big\{\bar{\boldsymbol{u}} \in \{0,1\}^{n_1} : \bar{u}_i = 0 \ \forall i \in \{1, \ldots, n_1 - n^-\} \cup \{n_1 - n^+ + 1, \ldots, n_1\}\big\} \setminus \{\boldsymbol{0}\}$$

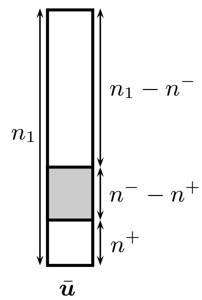

Fig. 16. A vector $\bar{\boldsymbol{u}}$ in the set $\bar{\mathcal{U}}'(n^-, n^+)$. White regions represent bits set to zero.

as illustrated in Fig. 16. We can then rewrite (58) as

$$\begin{aligned}\mu(B') \leq & \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_1, n_1 - \bar{R}_1)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} \mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) \\ & + \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_1, n_1 - \bar{R}_1)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} \mu\big(B'(\mathbf{0}, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) \\ & + \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} \mu\big(B'(\bar{\boldsymbol{u}}_0, \mathbf{0}, \bar{\boldsymbol{u}}_2)\big) \\ & + \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_1, n_1 - \bar{R}_1)} \mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \mathbf{0})\big).\end{aligned}$$

By (62), the set B' has measure zero whenever there is only a single minimizing $n(\bar{\boldsymbol{u}}_k)$. Together with the assumption $n_1 \geq n_0 \geq n_2$, this shows that we can restrict the lower boundaries of the sets $\bar{\mathcal{U}}'$ in the various sums. For example

$$\begin{aligned}& \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_1, n_1 - \bar{R}_1)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} \mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big) \\ & \qquad = \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_0, n_1 - \bar{R}_1)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} \mu\big(B'(\bar{\boldsymbol{u}}_0, \bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)\big),\end{aligned}$$

where we have changed $\bar{\mathcal{U}}'(n_1, n_1 - \bar{R}_1)$ to $\bar{\mathcal{U}}'(n_0, n_1 - \bar{R}_1)$, and similarly for the other three summations. Together with (61) this yields that

$$\begin{aligned}\mu(B') \leq & \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_0, n_1 - \bar{R}_1)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} 2^{\bar{R}_3 - n_1 + \min_k n(\bar{\boldsymbol{u}}_k)} \\ & + \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_2, n_1 - \bar{R}_1)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} 2^{\bar{R}_3 - n_1 + \min_k n(\bar{\boldsymbol{u}}_k)} \\ & + \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_2, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_2 \in \bar{\mathcal{U}}'(n_2, n_2 - \bar{R}_2)} 2^{\bar{R}_3 - n_1 + \min_k n(\bar{\boldsymbol{u}}_k)} \\ & + \sum_{\bar{\boldsymbol{u}}_0 \in \bar{\mathcal{U}}'(n_0, n_0 - \bar{R}_0)} \sum_{\bar{\boldsymbol{u}}_1 \in \bar{\mathcal{U}}'(n_0, n_1 - \bar{R}_1)} 2^{\bar{R}_3 - n_1 + \min_k n(\bar{\boldsymbol{u}}_k)}.\end{aligned}$$

We consider each of the four terms in turn.

For the first term, we have

$$\sum_{\bar{\boldsymbol{u}}_0\in\bar{\mathcal{U}}'(n_0,n_0-\bar{R}_0)}\ \sum_{\bar{\boldsymbol{u}}_1\in\bar{\mathcal{U}}'(n_0,n_1-\bar{R}_1)}\ \sum_{\bar{\boldsymbol{u}}_2\in\bar{\mathcal{U}}'(n_2,n_2-\bar{R}_2)} 2^{\bar{R}_3-n_1+\min_k n(\bar{\boldsymbol{u}}_k)}$$
$$= \sum_{i=n_1-n_0+1}^{n_1}\ \sum_{\bar{\boldsymbol{u}}_0\in\bar{\mathcal{U}}'(n_0,n_0-\bar{R}_0)}\ \sum_{\bar{\boldsymbol{u}}_1\in\bar{\mathcal{U}}'(n_0,n_1-\bar{R}_1)}\ \sum_{\bar{\boldsymbol{u}}_2\in\bar{\mathcal{U}}'(n_2,n_2-\bar{R}_2)} 2^{\bar{R}_3-n_1+i}\mathbb{1}_{\{\min_k n(\bar{\boldsymbol{u}}_k)=i\}}$$
$$\le \sum_{i=n_1-n_0+1}^{n_1}\ \sum_{\bar{\boldsymbol{u}}_0\in\bar{\mathcal{U}}'(n_1-i+1,n_0-\bar{R}_0)}\ \sum_{\bar{\boldsymbol{u}}_1\in\bar{\mathcal{U}}'(n_1-i+1,n_1-\bar{R}_1)}\ \sum_{\bar{\boldsymbol{u}}_2\in\bar{\mathcal{U}}'(n_2,n_2-\bar{R}_2)} 2^{\bar{R}_3-n_1+i}.$$

Using that

$$|\bar{\mathcal{U}}'(n^-,n^+)| = 2^{n^- - n^+}\mathbb{1}_{\{n^-\ge n^+\}} \le 2^{n^- - n^+},$$

the right-hand side can be further upper bounded by

$$\sum_{i=n_1-n_0+1}^{n_1} 2^{n_1-n_0+\bar{R}_0-i+1}\cdot 2^{\bar{R}_1-i+1}\cdot 2^{\bar{R}_2}\cdot 2^{\bar{R}_3-n_1+i}$$
$$\le 2^{\bar{R}_0+\bar{R}_1+\bar{R}_2+\bar{R}_3-n_0+2}\sum_{i=n_1-n_0+1}^{n_1} 2^{-i}$$
$$\le 2^{\bar{R}_0+\bar{R}_1+\bar{R}_2+\bar{R}_3-n_1+2}.$$

We can upper bound the remaining three terms in a similar fashion, yielding

$$\mu(B') \le 2^{\bar{R}_3+2}\big(2^{\bar{R}_0+\bar{R}_1+\bar{R}_2-n_1} + 2^{\bar{R}_1+\bar{R}_2-n_1} + 2^{\bar{R}_0+\bar{R}_2-n_0} + 2^{\bar{R}_0+\bar{R}_1-n_1}\big)$$
$$\le 16\cdot 2^{\bar{R}_3+\max\{\bar{R}_0+\bar{R}_1+\bar{R}_2-n_1,\bar{R}_0+\bar{R}_2-n_0\}}.$$

This shows that if

$$\bar{R}_1+\bar{R}_0+\bar{R}_2+\bar{R}_3 \le n_1-\log(16/\delta),$$
$$\bar{R}_0+\bar{R}_2+\bar{R}_3 \le n_0-\log(16/\delta),$$

and (in order to guarantee (56)) if

$$\bar{R}_2+\bar{R}_3 \le n_2,$$

then

$$\mu(B) \le \mu(B') \le \delta,$$

completing the proof of the lemma. ∎

B. Decoding Conditions for the Gaussian Channel

In this section, we analyze a "generic" receiver for the Gaussian case. To this end, we prove a variation of a well-known result from Diophantine approximation called Groshev's theorem (see, e.g., [23, Theorem 1.12]).

Define

$$\begin{aligned}\mathcal{U} \triangleq \big\{&u_k\in[-1,1]\ \forall k\{0,1,2\}:\\ &[u_0]_i = 0 \text{ for } i\in\{1,\dots,n_1-n_0\}\cup\{n_1-n_0+\bar{R}_0+1,\dots\},\\ &[u_1]_i = 0 \text{ for } i\in\{\bar{R}_1+1,\dots\},\\ &[u_2]_i = 0 \text{ for } i\in\{1,\dots,n_1-n_2\}\cup\{n_1-n_2+\bar{R}_2+1,\dots\},\\ &[u_3]_i = 0 \text{ for } i\in\{1,\dots,n_1-\bar{R}_3\}\cup\{n_1+1,\dots\}\big\}.\end{aligned}$$

$\mathcal{U}$ is the set of real numbers such that their binary expansion, when viewed as vectors of length n_1, is in the set $\bar{\mathcal{U}}$ as illustrated in Fig. 15 in Section VII-A. Here we assume that the binary expansion of u and $-u$ is identical.

Lemma 12. *Let $n_0, n_1, n_2 \in \mathbb{Z}_+$ such that $n_1 \geq n_0 \geq n_2$, and let $\bar{R}_0, \bar{R}_1, \bar{R}_2, \bar{R}_3 \in \mathbb{Z}_+$. Define the event*

$$B(u_0, u_1, u_2, u_3) \triangleq \big\{(g_0, g_1, g_2) \in (1, 4]^3 : |g_0 u_0 + g_1(u_1 + u_3) + g_2 u_2| \leq 2^{5-n_1}\big\},$$

and set

$$B \triangleq \bigcup_{(u_0,u_1,u_2,u_3)\in\mathcal{U}\setminus\{(0,0,0,0)\}} B(u_0, u_1, u_2, u_3).$$

If there exists $\delta \in (0, 1]$ satisfying

$$\begin{aligned} \bar{R}_1 + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_1 - 6 - \log(2184/\delta), \\ \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_0 - 6 - \log(2184/\delta), \\ \bar{R}_2 + \bar{R}_3 &\leq n_2 - 6, \end{aligned}$$

then

$$\mu(B) \leq \delta.$$

Lemma 12 is the equivalent for the Gaussian channel of Lemma 11 for the deterministic channel. Note that, except of the differences in the constants, the conditions on the rates in the two lemmas are identical.

We now prove Lemma 9 in Section VI-A using Lemma 12. Apply Lemma 12 with $\delta/2$ instead of δ and the same rate allocations as in the deterministic case, see Figs. 8 and 9 in Section IV-B. Let $\tilde{B}_m \subset (1, 4]^3$ be the collection of triples (g_{m0}, g_{m1}, g_{m2}) such that decoding is successful at receiver m. Define B_m as the collection of channel gains $(h_{mk}) \subset (1, 2]^{2\times 2}$ such that the corresponding (g_{mk}) are in $\tilde{B}_m$. Finally, let B denote the union of B_1 and B_2. Following the same arguments as in the proof of Lemma 7 from Lemma 11 presented in Section VII-A, it can be shown that if the decoding conditions in Lemma 9 are satisfied, then

$$\mu_3(\tilde{B}_m) \leq \delta/2$$

for $m \in \{1, 2\}$.

The next lemma allows us to transfer this statement about the products g_{mk} of channel gains to the corresponding statement about the original channel gains h_{mk}. For ease of notation, define

$$\begin{aligned} g_0 &\triangleq h_{11}h_{12}, \\ g_1 &\triangleq h_{11}h_{22}, \\ g_2 &\triangleq h_{12}h_{21}. \end{aligned}$$

Lemma 13. *Let $\tilde{B} \subseteq (1, 4]^3$ be a subset of channel gains (g_0, g_1, g_2) such that $\mu_3(\tilde{B}) \leq \delta$. Define*

$$B \triangleq \big\{(h_{mk}) \in (1, 2]^{2\times 2} : (g_0, g_1, g_2) \in \tilde{B}\big\}.$$

Then $\mu_4(B) \leq \delta$.

The proof of Lemma 13 is reported in Appendix E.

Applying Lemma 13, to the sets $\tilde{B}_1$ and $\tilde{B}_2$ corresponding to the outage events defined above, this implies that

$$\mu_4(B_m) \leq \delta/2.$$

Hence,

$$\mu_4(B) \leq \mu_4(B_1) + \mu_4(B_2) \leq \delta,$$

proving Lemma 9.

We continue with the proof of Lemma 12. Instead of directly analyzing the set B in the statement of Lemma 12, it will be convenient to work with an equivalent set. Note that $B(u_0, u_1, u_2, u_3)$ can be written as

$$B(u_0, u_1, u_2, u_3) = \{|g_0 2^{n_1} u_0 + g_1 2^{n_1}(u_1 + u_3) + g_2 2^{n_1} u_2| \leq 2^5\},$$

where we assume throughout that all sets are defined over $(g_0, g_1, g_2) \in (1, 4]^3$. By the definition of $\mathcal{U}$, we can decompose

$$\begin{aligned} 2^{n_1} u_0 &= A'_0 q_0, \\ 2^{n_1} u_1 &= A'_1 q_1, \\ 2^{n_1} u_2 &= A'_2 q_2, \\ 2^{n_1} u_3 &= q_3, \end{aligned}$$

with

$$A'_k \triangleq 2^{n_k - \bar{R}_k}$$

for $k \in \{0, 1, 2\}$ and

$$q_k \in \{-Q_k, -Q_k + 1, \ldots, Q_k - 1, Q_k\}$$

for $k \in \{0, 1, 2, 3\}$, where

$$Q_k \triangleq 2^{\bar{R}_k}.$$

Using this, we can further rewrite $B(u_0, u_1, u_2, u_3)$ using the triangle inequality as

$$\begin{aligned} B(u_0, u_1, u_2, u_3) &= \{|A'_0 g_0 q_0 + A'_1 g_1 q_1 + g_1 q_3 + A'_2 g_2 q_2| \leq 2^5\} \\ &\subseteq \{|A'_0 g_0 q_0 + A'_1 g_1 q_1 + A'_2 g_2 q_2| \leq 2^5 + 2^{\bar{R}_3 + 2}\} \\ &\subseteq \{|A'_0 g_0 q_0 + A'_1 g_1 q_1 + A'_2 g_2 q_2| \leq \beta'\} \\ &\triangleq B'(q_0, q_1, q_2), \end{aligned}$$

where we have defined

$$\beta' \triangleq 2^{\bar{R}_3 + 6}.$$

Setting

$$B' \triangleq \bigcup_{\substack{q_0, q_1, q_2 \in \mathbb{Z}: \\ (q_0, q_1, q_2) \neq \mathbf{0}, \\ |q_k| \leq Q_k \forall k}} B'(q_0, q_1, q_2),$$

we then have

$$\mu_3(B) \leq \mu_3(B').$$

The next lemma analyzes the set B' with $A'_0 = 1$.

Lemma 14. *Let* $\beta \in (0, 1]$, $A_1, A_2 \in \mathbb{N}$, *and* $Q_0, Q_1, Q_2 \in \mathbb{N}$. *Define the event*

$$B'(q_0, q_1, q_2) \triangleq \big\{(g_0, g_1, g_2) \in (1, 4]^3 : |g_0 q_0 + A_1 g_1 q_1 + A_2 g_2 q_2| < \beta\big\},$$

and set

$$B' \triangleq \bigcup_{\substack{q_0, q_1, q_2 \in \mathbb{Z}: \\ (q_0, q_1, q_2) \neq \mathbf{0}, \\ |q_k| \leq Q_k \forall k}} B'(q_0, q_1, q_2).$$

Then

$$\mu(B') \leq 168\beta\Big(2\min\Big\{Q_2, \frac{Q_0}{A_2}\Big\} + \min\Big\{Q_1\tilde{Q}_2, \frac{Q_0\tilde{Q}_2}{A_1}, \frac{A_2\tilde{Q}_2^2}{A_1}\Big\} \\ + 2\min\Big\{Q_1, \frac{Q_0}{A_1}\Big\} + \min\Big\{Q_2\tilde{Q}_1, \frac{Q_0\tilde{Q}_1}{A_2}, \frac{A_1\tilde{Q}_1^2}{A_2}\Big\}\Big)$$

with

$$\tilde{Q}_1 \triangleq \min\Big\{Q_1, 8\frac{\max\{Q_0, A_2Q_2\}}{A_1}\Big\},$$
$$\tilde{Q}_2 \triangleq \min\Big\{Q_2, 8\frac{\max\{Q_0, A_1Q_1\}}{A_2}\Big\}.$$

Remark: The special case of Lemma 14 with $A_1 = A_2 = 1$, $Q_0 = Q_1 = Q_2 = Q$, and $Q \to \infty$ corresponds to the (converse part of) Groshev's theorem, see, e.g., [23, Theorem 1.12]. Hence, Lemma 14 extends Groshev's theorem to asymmetric and non-asymptotic settings.

Before we present the proof of Lemma 14, we show how to prove Lemma 12 with the help of Lemma 14.

Proof of Lemma 12: We consider the cases $A'_0 \leq \min\{A'_1, A'_2\}$, $A'_1 \leq \min\{A'_0, A'_2\}$, and $A'_2 \leq \min\{A'_0, A'_1\}$, separately.

Assume first that $A'_0 \leq \min\{A'_1, A'_2\}$. Define

$$A_0 \triangleq 1,$$
$$A_1 \triangleq A'_1/A'_0 = 2^{\bar{R}_0-\bar{R}_1-n_0+n_1},$$
$$A_2 \triangleq A'_2/A'_0 = 2^{\bar{R}_0-\bar{R}_2-n_0+n_2},$$
$$\beta \triangleq \beta'/A'_0 = 2^{\bar{R}_0+\bar{R}_3-n_0+6}.$$

Note that $A_1, A_2 \in \mathbb{N}$, and that $\beta \in (0,1]$ if

$$\bar{R}_0 + \bar{R}_3 \leq n_0 - 6, \tag{63}$$

as required by the assumption in Lemma 14. The quantities $\tilde{Q}_1$ and $\tilde{Q}_2$ in Lemma 14 can be upper bounded as

$$\tilde{Q}_1 \leq 8\max\{Q_0, A_2Q_2\}/A_1 = 8Q_0/A_1,$$

since $n_0 \geq n_2$ implies that $Q_0 \geq A_2Q_2$, and

$$\tilde{Q}_2 \leq Q_2.$$

Applying Lemma 14 yields then that

$$\begin{aligned}
\mu(B) &\leq \mu(B') \\
&\leq 168\beta\Big(2Q_2 + \frac{A_2\tilde{Q}_2^2}{A_1} + 2\frac{Q_0}{A_1} + Q_2\tilde{Q}_1\Big) \\
&\leq 168\beta\Big(2Q_2 + \frac{A_2Q_2^2}{A_1} + 2\frac{Q_0}{A_1} + 8\frac{Q_0Q_2}{A_1}\Big) \\
&\leq 2184\beta\max\Big\{Q_2, \frac{A_2Q_2^2}{A_1}, \frac{Q_0}{A_1}, \frac{Q_0Q_2}{A_1}\Big\} \\
&= 2184\beta\max\Big\{Q_2, \frac{Q_0Q_2}{A_1}\Big\},
\end{aligned}$$

where we have used that $Q_2 \geq 1$ and that $A_2 Q_2 \leq Q_0$ implying

$$\frac{A_2 Q_2^2}{A_1} \leq \frac{Q_0 Q_2}{A_1}.$$

Substituting the definitions of β, A_k, and Q_k, yields that

$$\mu(B) \leq 2184 \cdot 2^{\bar{R}_0+\bar{R}_3-n_0+6} \max\left\{2^{\bar{R}_2}, 2^{\bar{R}_1+\bar{R}_2+n_0-n_1}\right\}.$$

Together with (63), this shows that if

$$\begin{aligned}\bar{R}_1+\bar{R}_0+\bar{R}_2+\bar{R}_3 &\leq n_1-6-\log(2184/\delta),\\ \bar{R}_0+\bar{R}_2+\bar{R}_3 &\leq n_0-6-\log(2184/\delta),\\ \bar{R}_0+\bar{R}_3 &\leq n_0-6,\end{aligned}$$

then

$$\mu(B) \leq \delta.$$

Since $\delta \in (0,1]$ and $\bar{R}_2 \geq 0$, the third condition is redundant and can be removed, showing the result in Lemma 12. We point out that the third condition in Lemma 12 is not active if $A'_0 \leq \min\{A'_1, A'_2\}$. This is consistent with it not appearing in the derivation here.

Assume next that $A'_1 \leq \min\{A'_0, A'_2\}$. Define

$$\begin{aligned}A_0 &\triangleq A'_0/A'_1 = 2^{-\bar{R}_0+\bar{R}_1+n_0-n_1},\\ A_1 &\triangleq 1,\\ A_2 &\triangleq A'_2/A'_1 = 2^{\bar{R}_1-\bar{R}_2-n_1+n_2},\\ \beta &\triangleq \beta'/A'_1 = 2^{\bar{R}_1+\bar{R}_3-n_1+6}.\end{aligned}$$

Note that $A_0, A_2 \in \mathbb{N}$, and that $\beta \in (0,1]$ if

$$\bar{R}_1+\bar{R}_3 \leq n_1-6. \tag{64}$$

We can hence apply Lemma 14 by appropriately relabeling indices (i.e., by swapping indices 0 and 1). The quantities $\tilde{Q}_0$ and $\tilde{Q}_2$ can be upper bounded as

$$\tilde{Q}_0 \triangleq \min\left\{Q_0, 8\frac{\max\{Q_1, A_2Q_2\}}{A_0}\right\} \leq Q_0,$$

and

$$\tilde{Q}_2 \triangleq \min\left\{Q_2, 8\frac{\max\{Q_1, A_0Q_0\}}{A_2}\right\} \leq Q_2.$$

Applying Lemma 14 yields then that

$$\begin{aligned}\mu(B) &\leq \mu(B')\\ &\leq 168\beta\Big(2Q_2+Q_0\tilde{Q}_2+2Q_0+Q_2\tilde{Q}_0\Big)\\ &\leq 168\beta\Big(2Q_2+Q_0Q_2+2Q_0+Q_2Q_0\Big)\\ &\leq 1008\beta Q_0Q_2.\end{aligned}$$

Substituting the definitions of β, A_k, and Q_k, yields that

$$\mu(B) \leq 1008 \cdot 2^{\bar{R}_0+\bar{R}_1+\bar{R}_2+\bar{R}_3+6-n_1}.$$

Together with (64), this shows that if

$$\bar{R}_1 + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 \leq n_1 - 6 - \log(1008/\delta),$$
$$\bar{R}_1 + \bar{R}_3 \leq n_1 - 6,$$

then

$$\mu(B) \leq \delta.$$

Since $\delta \in (0,1]$ and $\bar{R}_0, \bar{R}_2 \geq 0$, the second condition is redundant and can be removed, showing the result in Lemma 12. As can be verified, the second and third conditions in Lemma 12 are not active when $A'_1 \leq \min\{A'_0, A'_2\}$, consistent with them not appearing in the derivation here.

Finally, assume that $A'_2 \leq \min\{A'_0, A'_1\}$. Define

$$\begin{aligned} A_0 &\triangleq A'_0/A'_2 = 2^{-\bar{R}_0+\bar{R}_2+n_0-n_2}, \\ A_1 &\triangleq A'_1/A'_2 = 2^{-\bar{R}_1+\bar{R}_2+n_1-n_2}, \\ A_2 &\triangleq 1, \\ \beta &\triangleq \beta'/A'_2 = 2^{\bar{R}_2+\bar{R}_3-n_2+6}. \end{aligned}$$

Note that $A_0, A_1 \in \mathbb{N}$, and that $\beta \in (0,1]$ if

$$\bar{R}_2 + \bar{R}_3 \leq n_2 - 6. \tag{65}$$

We can hence apply Lemma 14 by relabeling indices as before (this time by swapping indices 0 and 2). The quantities $\tilde{Q}_0$ and $\tilde{Q}_1$ can be upper bounded as

$$\tilde{Q}_0 \triangleq \min\left\{Q_0, 8\frac{\max\{Q_2, A_1 Q_1\}}{A_0}\right\} \leq Q_0,$$

and

$$\begin{aligned} \tilde{Q}_1 &\triangleq \min\left\{Q_1, 8\frac{\max\{Q_2, A_0 Q_0\}}{A_1}\right\} \\ &= 8\frac{\max\{Q_2, A_0 Q_0\}}{A_1} \\ &= 8\frac{A_0 Q_0}{A_1} \end{aligned}$$

since $n_0 \geq n_2$ implies $A_0 Q_0 \geq Q_2$.

Applying Lemma 14 yields then that

$$\begin{aligned} \mu(B) &\leq \mu(B') \\ &\leq 168\beta\Big(2\frac{Q_2}{A_0} + \frac{Q_2\tilde{Q}_0}{A_1} + 2\frac{Q_2}{A_1} + \frac{Q_2\tilde{Q}_1}{A_0}\Big) \\ &\leq 168\beta\Big(2\frac{Q_2}{A_0} + \frac{Q_2 Q_0}{A_1} + 2\frac{Q_2}{A_1} + 8\frac{Q_2 Q_0}{A_1}\Big) \\ &\leq 2184\beta\max\left\{\frac{Q_2}{A_0}, \frac{Q_2 Q_0}{A_1}\right\}. \end{aligned}$$

Substituting the definitions of β, A_k, and Q_k, yields that

$$\mu(B) \leq 2184 \cdot 2^{\bar{R}_2+\bar{R}_3+6-n_2} \max\left\{2^{\bar{R}_0+n_2-n_0}, 2^{\bar{R}_0+\bar{R}_1+n_2-n_1}\right\}.$$

Together with (65), this shows that if

$$\begin{aligned}\bar{R}_1 + \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_1 - 6 - \log(2184/\delta),\\ \bar{R}_0 + \bar{R}_2 + \bar{R}_3 &\leq n_0 - 6 - \log(2184/\delta),\\ \bar{R}_2 + \bar{R}_3 &\leq n_2 - 6,\end{aligned}$$

then

$$\mu(B) \leq \delta,$$

showing the result in Lemma 12. It can be verified that, unlike in the other two cases, all three conditions in Lemma 12 can be active when $A'_2 \leq \min\{A'_0, A'_1\}$. This is again consistent with the derivation here. This proves Lemma 12. ∎

It remains to prove Lemma 14. The proof builds on an argument in [24].

Proof of Lemma 14: Define

$$B'(q_1, q_2) \triangleq \bigcup_{\substack{q_0 \in \mathbb{Z}:\\ |q_0| \leq Q_0}} B'(q_0, q_1, q_2)$$

for $(q_1, q_2) \neq (0, 0)$, and

$$B'(0, 0) \triangleq \bigcup_{\substack{q_0 \in \mathbb{Z} \setminus \{0\}:\\ |q_0| \leq Q_0}} B'(q_0, q_1, q_2).$$

For $g_0 \in (1, 4]$, set

$$B'_{g_0}(q_1, q_2) \triangleq \big\{(g_1, g_2) \in (1, 4]^2 : (g_0, g_1, g_2) \in B'(q_1, q_2)\big\}.$$

Observe that $B'_{g_0}(q_1, q_2)$ is a subset of $\mathbb{R}^2$ and that

$$\mu_3(B'(q_1, q_2)) = \int_{g_0=1}^{4} \mu_2(B'_{g_0}(q_1, q_2)) dg_0.$$

We treat the cases $A_1|q_1| \leq A_2|q_2|$ and $A_1|q_1| \geq A_2|q_2|$ separately. Assume first $A_1|q_1| \leq A_2|q_2|$ and $q_2 \neq 0$. If

$$A_2|q_2| \geq 8 \max\{Q_0, A_1 Q_1\} + 1,$$

then

$$\begin{aligned}|g_0 q_0 + A_1 g_1 q_1 + A_2 g_2 q_2| &\geq A_2 g_2 |q_2| - A_1 g_1 |q_1| - g_0 |q_0|\\ &\geq A_2|q_2| - 4A_1 Q_1 - 4Q_0\\ &\geq 1\\ &\geq \beta,\end{aligned}$$

where we have used that $\beta \leq 1$. Hence, $\mu_2(B'_{g_0}(q_1, q_2)) = 0$. We can therefore assume without loss of generality that

$$A_2|q_2| \leq 8 \max\{Q_0, A_1 Q_1\}.$$

By a similar argument, we can assume that

$$A_2|q_2| \leq 4Q_0$$

whenever $q_1 = 0$. The set $B'_{g_0}(q_1, q_2)$ consists of at most

$$\min\{3Q_0, 7A_2|q_2|\}$$

strips of slope $-A_1q_1/(A_2q_2)$ and width $2\beta/(A_2|q_2|)$ in the g_2 direction, including several partial strips (see Fig. 17). The area of this set is at most

$$\begin{aligned}\mu_2(B'_{g_0}(q_1,q_2)) &\leq \frac{2\beta}{A_2|q_2|}\min\{3Q_0, 7A_2|q_2|\} \\ &\leq 14\beta\min\left\{\frac{Q_0}{A_2|q_2|}, 1\right\}. \end{aligned} \tag{66}$$

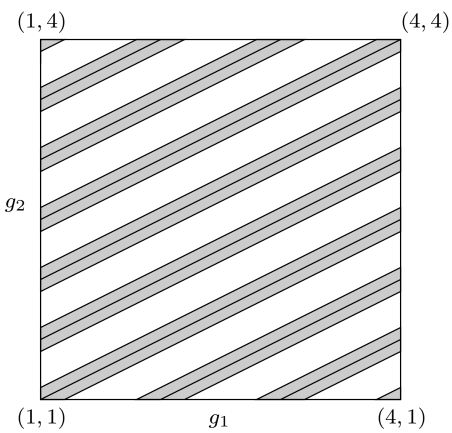

Fig. 17. Illustration of the set $B'_1(1,-2) \subseteq (1,4]^2$ with $g_0 = 1, q_1 = 1, q_2 = -2$, $A_1 = A_2 = 1$, and $\beta = 0.2$. In the figure, we assume that $Q_0 \gg A_2|q_2|$. The set consists of $10 \leq 7A_2|q_2| = 14$ strips of slope $1/2 = -A_1q_1/(A_2q_2)$.

We now consider the case $A_1|q_1| \geq A_2|q_2|$ and $q_1 \neq 0$. As before, we can assume without loss of generality that

$$A_1|q_1| \leq 8\max\{Q_0, A_2Q_2\}$$

for all q_2, and that

$$A_1|q_1| \leq 4Q_0$$

whenever $q_2 = 0$. By the same analysis as in the last paragraph, we obtain that

$$\mu_2(B'_{g_0}(q_1,q_2)) \leq 14\beta\min\left\{\frac{Q_0}{A_1|q_1|}, 1\right\}. \tag{67}$$

Finally, when $(q_1, q_2) = \mathbf{0}$ and $q_0 \neq 0$, then $g_0|q_0| \geq 1 \geq \beta$, and hence

$$\mu_2(B'_{g_0}(0,0)) = 0. \tag{68}$$

We can upper bound

$$\begin{aligned}\mu_3(B') &= \mu_3\Bigg(\cup_{\substack{q_1\in\mathbb{Z}:\\|q_1|\leq Q_1}} \cup_{\substack{q_2\in\mathbb{Z}:\\|q_2|\leq Q_2}} \cup_{\substack{q_0\in\mathbb{Z}:|q_0|\leq Q_0\\(q_0,q_1,q_2)\neq\mathbf{0}}} B'(q_0,q_1,q_2)\Bigg) \\ &\leq \sum_{\substack{q_1\in\mathbb{Z}:\\|q_1|\leq Q_1}}\sum_{\substack{q_2\in\mathbb{Z}:\\|q_2|\leq Q_2}} \mu_3(B'(q_1,q_2)) \\ &= \sum_{\substack{q_1\in\mathbb{Z}:\\|q_1|\leq Q_1}}\sum_{\substack{q_2\in\mathbb{Z}:\\|q_2|\leq Q_2}} \int_{g_0=1}^{4} \mu_2(B'_{g_0}(q_1,q_2))dg_0 \\ &= \sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\leq Q_2\\A_2|q_2|\leq 4Q_0}} \int_{g_0=1}^{4} \mu_2(B'_{g_0}(0,q_2))dg_0 \end{aligned}$$

$$+\sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2\\ A_2|q_2|\le 8\max\{Q_0,A_1Q_1\}}}\ \sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le A_2|q_2|}}\int_{g_0=1}^{4}\mu_2(B'_{g_0}(q_1,q_2))dg_0$$

$$+\sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le 4Q_0}}\int_{g_0=1}^{4}\mu_2(B'_{g_0}(q_1,0))dg_0$$

$$+\sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le 8\max\{Q_0,A_2Q_2\}}}\ \sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2\\ A_2|q_2|\le A_1|q_1|}}\int_{g_0=1}^{4}\mu_2(B'_{g_0}(q_1,q_2))dg_0.$$

Combined with (66), (67), and (68), this yields

$$\begin{aligned}\mu_3(B') \le &\sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2\\ A_2|q_2|\le 4Q_0}} 42\beta\min\left\{\frac{Q_0}{A_2|q_2|},1\right\}\\
&+\sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2\\ A_2|q_2|\le 8\max\{Q_0,A_1Q_1\}}}\ \sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le A_2|q_2|}} 42\beta\min\left\{\frac{Q_0}{A_2|q_2|},1\right\}\\
&+\sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le 4Q_0}} 42\beta\min\left\{\frac{Q_0}{A_1|q_1|},1\right\}\\
&+\sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le 8\max\{Q_0,A_2Q_2\}}}\ \sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2\\ A_2|q_2|\le A_1|q_1|}} 42\beta\min\left\{\frac{Q_0}{A_1|q_1|},1\right\}. \end{aligned}\tag{69}$$

We now upper bound the four terms in the right-hand side of (69).

For the first term, observe that

$$\begin{aligned}\left|\left\{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2, A_2|q_2|\le 4Q_0\right\}\right| &\le 2\min\left\{Q_2, 4\frac{Q_0}{A_2}\right\}\\
&\le 8\min\left\{Q_2,\frac{Q_0}{A_2}\right\},\end{aligned}$$

so that

$$\sum_{\substack{q_2\in\mathbb{Z}\setminus\{0\}:|q_2|\le Q_2\\ A_2|q_2|\le 4Q_0}} 42\beta\min\left\{\frac{Q_0}{A_2|q_2|},1\right\}\le 336\beta\min\left\{Q_2,\frac{Q_0}{A_2}\right\}.\tag{70}$$

For the second term in (69), observe that

$$\left|\left\{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1, A_1|q_1|\le A_2|q_2|\right\}\right|\le 2\min\left\{Q_1,\frac{A_2|q_2|}{A_1}\right\},$$

and hence

$$\begin{aligned}\sum_{\substack{q_1\in\mathbb{Z}\setminus\{0\}:|q_1|\le Q_1\\ A_1|q_1|\le A_2|q_2|}}\min\left\{\frac{Q_0}{A_2|q_2|},1\right\} &\le 2\min\left\{Q_1,\frac{A_2|q_2|}{A_1}\right\}\min\left\{\frac{Q_0}{A_2|q_2|},1\right\}\\
&\le 2\min\left\{\frac{Q_0Q_1}{A_2|q_2|},Q_1,\frac{Q_0}{A_1},\frac{A_2|q_2|}{A_1}\right\}\\
&\le 2\min\left\{Q_1,\frac{Q_0}{A_1},\frac{A_2|q_2|}{A_1}\right\}.\end{aligned}$$

Moreover,

$$\left\{q_2 \in \mathbb{Z} \setminus \{0\} : |q_2| \le Q_2, A_2|q_2| \le 8 \max\{Q_0, A_1Q_1\}\right\} = \left\{q_2 \in \mathbb{Z} \setminus \{0\} : |q_2| \le \tilde{Q}_2\right\}$$

with

$$\tilde{Q}_2 \triangleq \min\left\{Q_2, 8\frac{\max\{Q_0, A_1Q_1\}}{A_2}\right\}.$$

Using these two facts, we can upper bound

$$\begin{aligned}
\sum_{\substack{q_2 \in \mathbb{Z}\setminus\{0\}:|q_2|\le Q_2 \\ A_2|q_2|\le 8\max\{Q_0,A_1Q_1\}}} \; \sum_{\substack{q_1 \in \mathbb{Z}\setminus\{0\}:|q_1|\le Q_1 \\ A_1|q_1|\le A_2|q_2|}} & 42\beta \min\left\{\frac{Q_0}{A_2|q_2|}, 1\right\} \\
& \le 84\beta \sum_{q_2 \in \mathbb{Z}\setminus\{0\}:|q_2|\le \tilde{Q}_2} \min\left\{Q_1, \frac{Q_0}{A_1}, \frac{A_2|q_2|}{A_1}\right\} \\
& \le 84\beta \sum_{q_2 \in \mathbb{Z}\setminus\{0\}:|q_2|\le \tilde{Q}_2} \min\left\{Q_1, \frac{Q_0}{A_1}, \frac{A_2\tilde{Q}_2}{A_1}\right\} \\
& \le 168\beta \min\left\{Q_1\tilde{Q}_2, \frac{Q_0\tilde{Q}_2}{A_1}, \frac{A_2\tilde{Q}_2^2}{A_1}\right\}. \qquad (71)
\end{aligned}$$

Similarly, for the third term in (69),

$$\sum_{\substack{q_1 \in \mathbb{Z}\setminus\{0\}:|q_1|\le Q_1 \\ A_1|q_1|\le 4Q_0}} 42\beta \min\left\{\frac{Q_0}{A_1|q_1|}, 1\right\} \le 336\beta \min\left\{Q_1, \frac{Q_0}{A_1}\right\}, \qquad (72)$$

and for the fourth term

$$\sum_{\substack{q_1 \in \mathbb{Z}\setminus\{0\}:|q_1|\le Q_1 \\ A_1|q_1|\le 8\max\{Q_0,A_2Q_2\}}} \; \sum_{\substack{q_2 \in \mathbb{Z}\setminus\{0\}:|q_2|\le Q_2 \\ A_2|q_2|\le A_1|q_1|}} 42\beta \min\left\{\frac{Q_0}{A_1|q_1|}, 1\right\} \le 168\beta \min\left\{Q_2\tilde{Q}_1, \frac{Q_0\tilde{Q}_1}{A_2}, \frac{A_1\tilde{Q}_1^2}{A_2}\right\} \qquad (73)$$

with the definition

$$\tilde{Q}_1 \triangleq \min\left\{Q_1, 8\frac{\max\{Q_0, A_2Q_2\}}{A_1}\right\}.$$

Substituting (70)–(73) into (69) yields

$$\begin{aligned}
\mu_3(B') \le \beta\Bigg(& 336 \min\left\{Q_2, \frac{Q_0}{A_2}\right\} + 168 \min\left\{Q_1\tilde{Q}_2, \frac{Q_0\tilde{Q}_2}{A_1}, \frac{A_2\tilde{Q}_2^2}{A_1}\right\} \\
& + 336 \min\left\{Q_1, \frac{Q_0}{A_1}\right\} + 168 \min\left\{Q_2\tilde{Q}_1, \frac{Q_0\tilde{Q}_1}{A_2}, \frac{A_1\tilde{Q}_1^2}{A_2}\right\}\Bigg),
\end{aligned}$$

completing the proof. ■

VIII. Conclusion

In this paper, we derived a constant-gap capacity approximation for the Gaussian X-channel. This derivation was aided by a novel deterministic channel model used to approximate the Gaussian channel. In the proposed deterministic channel model, the actions of the channel are described by a lower-triangular Toeplitz matrix with coefficients determined by the bits in the binary expansion of the corresponding channel gain in the original Gaussian problem. This is in contrast to traditional deterministic models, in which the actions of the channel are only dependent on the single most-important bit of the channel gain

in the original Gaussian problem. Preserving this dependence on the fine structure of the Gaussian channel gains turned out to be crucial to approximate the Gaussian X-channel by a deterministic channel model.

Throughout this paper, we were only interested in obtaining a constant-gap capacity approximation. Less emphasis was placed on the actual value of that constant. For a meaningful capacity approximation at smaller values of SNR, this constant needs to be optimized. More sophisticated lattice codes (as opposed to the ones over the simple integer lattice used in this paper) could be employed for this purpose [25]. Furthermore, all the results in this paper were derived for all channel gains outside an arbitrarily small "outage" set. Analyzing the behavior of capacity for channel gains that are inside this outage set is hence of interest. An approach similar to the one in [26] could perhaps be utilized to this end.

Finally, the analysis in this paper focused on the Gaussian X-channel as the simplest communication network in which interference alignment seems necessary. The hope is that the tools developed in this paper can be used to help with the analysis of more general networks requiring interference alignment. Ultimately, the goal should be to move from degrees-of-freedom capacity approximations to stronger constant-gap capacity approximations.

Appendix A
Verification of Decoding Conditions

This appendix verifies that the rate allocation in Section V-A for the deterministic X-channel satisfies the decoding conditions (19) and (20) in Lemma 7.

Case I ($n_{12}+n_{21}\leq n_{11}$): Recall

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11}-n_{21},\\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22}-n_{12},\\
\bar{R}_{22}^{\mathsf{C}} &\triangleq \bar{R}_{11}^{\mathsf{C}} \triangleq \bar{R}_{12} \triangleq \bar{R}_{21} \triangleq 0.
\end{aligned}$$

This choice of rates satisfies (19a) and (20a). Since these are the only two relevant conditions in this case, this shows that both receivers can recover the desired messages.

Case II ($n_{11}<n_{12}+n_{21}\leq n_{22}$): Recall

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11}-n_{21},\\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22}-n_{12},\\
\bar{R}_{22}^{\mathsf{C}} &\triangleq n_{12}-\bar{R}_{11}^{\mathsf{P}},\\
\bar{R}_{11}^{\mathsf{C}} &\triangleq \bar{R}_{12} \triangleq \bar{R}_{21} \triangleq 0.
\end{aligned}$$

At receiver one, (19a) and (19b) are satisfied since

$$\bar{R}_{22}^{\mathsf{C}}+\bar{R}_{11}^{\mathsf{P}} = n_{12} \leq n_{11}.$$

Condition (19c) needs not to be checked here. At receiver two, (20a) is satisfied since

$$\bar{R}_{22}^{\mathsf{C}}+\bar{R}_{22}^{\mathsf{P}} = n_{22}+n_{21}-n_{11} \leq n_{22}.$$

Conditions (20b) and (20c) need not to be checked here. Hence both receivers can decode successfully.

Case III ($n_{22}<n_{12}+n_{21}\leq n_{11}+\frac{1}{2}n_{22}$): Recall

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11}-n_{21},\\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22}-n_{12},\\
\bar{R}_{12} &\triangleq (n_{12}+2n_{21}-n_{11}-n_{22})^{+},\\
\bar{R}_{21} &\triangleq (n_{21}+2n_{12}-n_{11}-n_{22})^{+},\\
\bar{R}_{11}^{\mathsf{C}} &\triangleq n_{21}-\bar{R}_{22}^{\mathsf{P}}-\bar{R}_{21},\\
\bar{R}_{22}^{\mathsf{C}} &\triangleq n_{12}-\bar{R}_{11}^{\mathsf{P}}-\bar{R}_{12}.
\end{aligned}$$

To check the decoding conditions (19) and (20), we first argue that

$$\max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} = \bar{R}_{22}^{\mathsf{C}}, \tag{74a}$$
$$\max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} = \bar{R}_{11}^{\mathsf{C}}. \tag{74b}$$

The first equality trivially holds if $\bar{R}_{21} = 0$. Assuming then that $\bar{R}_{21} > 0$, we have

$$\bar{R}_{22}^{\mathsf{C}} - \bar{R}_{21} = n_{22} - n_{12} - \bar{R}_{12}.$$

If $\bar{R}_{12} = 0$, then this is nonnegative. Assuming then $\bar{R}_{12} > 0$, we obtain

$$\bar{R}_{22}^{\mathsf{C}} - \bar{R}_{21} = 2n_{22} + n_{11} - 2(n_{12} + n_{21}) \geq 0,$$

where we have used that $n_{22} \geq n_{11}$ and that $n_{11} + \frac{1}{2}n_{22} \geq n_{12} + n_{21}$. This proves (74a). Using a similar argument, it can be shown that $\bar{R}_{11}^{\mathsf{C}} - \bar{R}_{12} \geq 0$, proving (74b). To check the decoding conditions (19) at receiver one, observe now that

$$\bar{R}_{22}^{\mathsf{C}} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} = n_{12},$$

satisfying (19b). Moreover,

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{C}} + \bar{R}_{22}^{\mathsf{C}} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} &= \bar{R}_{11}^{\mathsf{C}} + n_{12} \\
&= 2n_{12} + n_{21} - n_{22} - \bar{R}_{21} \\
&\leq n_{11}
\end{aligned}$$

satisfying (19a). Finally, if $\bar{R}_{12} > 0$, then

$$\bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} = n_{12} + n_{21} - n_{22},$$

satisfying (19c), and if $\bar{R}_{12} = 0$, (19c) is irrelevant. Using a similar argument, it can be shown that the decoding conditions (20) at receiver two hold.

Case IV ($n_{11} + \frac{1}{2}n_{22} < n_{12} + n_{21} \leq \frac{3}{2}n_{22}$): Recall

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11} - n_{21}, \\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22} - n_{12}, \\
\bar{R}_{21} &\triangleq \left\lfloor n_{12} - \tfrac{1}{2}n_{22} \right\rfloor, \\
\bar{R}_{12} &\triangleq \bar{R}_{11}^{\mathsf{C}} \triangleq \left\lfloor n_{21} - \tfrac{1}{2}n_{22} \right\rfloor, \\
\bar{R}_{22}^{\mathsf{C}} &\triangleq n_{22} - n_{21}.
\end{aligned}$$

To check the decoding conditions, note first that

$$\max\{\bar{R}_{21}, \bar{R}_{22}^{\mathsf{C}}\} = \bar{R}_{22}^{\mathsf{C}},$$
$$\max\{\bar{R}_{12}, \bar{R}_{11}^{\mathsf{C}}\} = \bar{R}_{11}^{\mathsf{C}},$$

since

$$\tfrac{3}{2}n_{22} \geq n_{12} + n_{21}$$

by assumption. For receiver one, we then have

$$\bar{R}_{11}^{\mathsf{C}} + \bar{R}_{22}^{\mathsf{C}} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11},$$

satisfying (19a). Moreover,

$$\bar{R}_{22}^{\mathsf{C}} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11} + \tfrac{1}{2}n_{22} - n_{21} \leq n_{12},$$

where we have used $n_{11} + \frac{1}{2}n_{22} < n_{12} + n_{21}$. Hence (19b) is satisfied. Finally,

$$\bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11} - \tfrac{1}{2}n_{22} \leq n_{12} + n_{21} - n_{22},$$

where we have again used $n_{11} + \frac{1}{2}n_{22} < n_{12} + n_{21}$. Hence (19c) is satisfied. Together, this shows that decoding is successful at receiver one. At receiver two, we have

$$\bar{R}_{22}^{\mathsf{C}} + \bar{R}_{11}^{\mathsf{C}} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}} \leq n_{22},$$

satisfying (20a), and

$$\bar{R}_{11}^{\mathsf{C}} + \bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}} \leq n_{21},$$

satisfying (20b). Finally,

$$\bar{R}_{21} + \bar{R}_{22}^{\mathsf{P}} \leq \tfrac{1}{2}n_{22} \leq n_{12} + n_{21} - n_{11},$$

where we have used $n_{11} + \frac{1}{2}n_{22} < n_{12} + n_{21}$. Hence (20c) is satisfied. Together, this shows that decoding is successful at receiver two.

Case V ($\frac{3}{2}n_{22} < n_{12} + n_{21} \leq n_{11} + n_{22}$): Recall

$$\begin{aligned}
\bar{R}_{11}^{\mathsf{P}} &\triangleq n_{11} - n_{21}, \\
\bar{R}_{22}^{\mathsf{P}} &\triangleq n_{22} - n_{12}, \\
\bar{R}_{12} &\triangleq \bar{R}_{11}^{\mathsf{C}} \triangleq \left\lfloor \tfrac{2}{3}n_{21} - \tfrac{1}{3}n_{12} \right\rfloor, \\
\bar{R}_{21} &\triangleq \bar{R}_{22}^{\mathsf{C}} \triangleq \left\lfloor \tfrac{2}{3}n_{12} - \tfrac{1}{3}n_{21} \right\rfloor.
\end{aligned}$$

For decoding at receiver one, we need to verify the decoding conditions (19). We have

$$\bar{R}_{11}^{\mathsf{C}} + \bar{R}_{22}^{\mathsf{C}} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11},$$

satisfying (19a). Moreover,

$$\bar{R}_{22}^{\mathsf{C}} + \bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} \leq n_{11} - \tfrac{1}{3}(2n_{21} - n_{12}) \leq n_{12},$$

where we have used (27). This satisfies (19b). Finally

$$\begin{aligned}
\bar{R}_{12} + \bar{R}_{11}^{\mathsf{P}} &\leq n_{11} - \tfrac{1}{3}(n_{12} + n_{21}) \\
&= n_{11} + n_{12} + n_{21} - \tfrac{4}{3}(n_{12} + n_{21}) \\
&\leq n_{11} + n_{12} + n_{21} - 2n_{22} \\
&\leq n_{12} + n_{21} - n_{22},
\end{aligned}$$

where we have used that $n_{12} + n_{21} \geq \frac{3}{2}n_{22}$. Hence (19c) is satisfied. A similar argument using (26) instead of (27), shows that the decoding conditions (20) at receiver two hold. Hence decoding is successful at both receivers.

Appendix B
Proof of Lemma 8 in Section V-B

Throughout this proof, we make use of the fact that, for the (modulated) deterministic X-channel (17), the definition of capacity imposes that

$$\bar{\boldsymbol{u}}_{mk}^{(T)} \triangleq (\bar{\boldsymbol{u}}_{mk}[t])_{t=1}^{T}$$

is only a function of w_{mk}.

We start with (29a). Define $\bar{\boldsymbol{s}}_{12}$ as the contribution of the second transmitter at the first receiver, i.e.,

$$\bar{\boldsymbol{s}}_{12} \triangleq \bar{\boldsymbol{G}}_{12} \begin{pmatrix} \boldsymbol{0} \\ \bar{\boldsymbol{u}}_{12}^{\mathsf{C}} \end{pmatrix} \oplus \bar{\boldsymbol{G}}_{10} \begin{pmatrix} \boldsymbol{0} \\ \bar{\boldsymbol{u}}_{22}^{\mathsf{C}} \end{pmatrix}.$$

Let $\bar{\boldsymbol{s}}_{22}$ denote the contribution of the second transmitter at the second receiver, i.e.,

$$\bar{\boldsymbol{s}}_{22} \triangleq \bar{\boldsymbol{G}}_{22}\bar{\boldsymbol{u}}_{22} \oplus \bar{\boldsymbol{G}}_{20}\bar{\boldsymbol{u}}_{12}.$$

Similarly, we define $\bar{\boldsymbol{s}}_{11}$ and $\bar{\boldsymbol{s}}_{21}$ as the contributions of the first transmitter at the first and second receivers, respectively. With this, we can rewrite the received vector at receiver m as

$$\bar{\boldsymbol{y}}_m = \bar{\boldsymbol{s}}_{m1} \oplus \bar{\boldsymbol{s}}_{m2}.$$

For block length T, we have,

$$\begin{aligned} T(\bar{R}_{22} - \varepsilon) &\leq I\big(w_{22}; \bar{\boldsymbol{y}}_2^{(T)}\big) \\ &\leq I\big(w_{22}; \bar{\boldsymbol{y}}_2^{(T)}, \bar{\boldsymbol{s}}_{12}^{(T)}, \bar{\boldsymbol{u}}_{11}^{(T)}, \bar{\boldsymbol{u}}_{21}^{(T)}, w_{12}\big) \\ &= I\big(w_{22}; \bar{\boldsymbol{y}}_2^{(T)}, \bar{\boldsymbol{s}}_{12}^{(T)} \bigm| \bar{\boldsymbol{u}}_{11}^{(T)}, \bar{\boldsymbol{u}}_{21}^{(T)}, w_{12}\big) \\ &= I\big(w_{22}; \bar{\boldsymbol{s}}_{22}^{(T)}, \bar{\boldsymbol{s}}_{12}^{(T)} \bigm| w_{12}\big) \\ &= I\big(w_{22}; \bar{\boldsymbol{s}}_{12}^{(T)} \bigm| w_{12}\big) + I\big(w_{22}; \bar{\boldsymbol{s}}_{22}^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big) \\ &\leq H\big(\bar{\boldsymbol{s}}_{12}^{(T)} \bigm| w_{12}\big) + H\big(\bar{\boldsymbol{s}}_{22}^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big), \end{aligned} \tag{75}$$

where the first step follows from Fano's inequality. In addition, using again Fano's inequality,

$$\begin{aligned} T(\bar{R}_{11} + \bar{R}_{12} - \varepsilon) &\leq I\big(w_{11}, w_{12}; \bar{\boldsymbol{y}}_1^{(T)}\big) \\ &\leq I\big(w_{11}, w_{12}, w_{21}; \bar{\boldsymbol{y}}_1^{(T)}\big) \\ &= H\big(\bar{\boldsymbol{y}}_1^{(T)}\big) - H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| w_{11}, w_{12}, w_{21}\big) \\ &= H\big(\bar{\boldsymbol{y}}_1^{(T)}\big) - H\big(\bar{\boldsymbol{s}}_{12}^{(T)} \bigm| w_{12}\big). \end{aligned} \tag{76}$$

Adding (75) and (76) yields

$$T(\bar{R}_{11} + \bar{R}_{12} + \bar{R}_{22} - 2\varepsilon) \leq H\big(\bar{\boldsymbol{y}}_1^{(T)}\big) + H\big(\bar{\boldsymbol{s}}_{22}^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big).$$

For the first term on the right-hand side, we have

$$H\big(\bar{\boldsymbol{y}}_1^{(T)}\big) \leq T \max\{n_{11}, n_{12}\}.$$

For the second term, recall that $\bar{\boldsymbol{u}}_{12}^{(T)}$ is a function of only w_{12}, and hence

$$\begin{aligned} H\big(\bar{\boldsymbol{s}}_{22}^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big) &\leq H\big(\bar{\boldsymbol{s}}_{22}^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, \bar{\boldsymbol{u}}_{12}^{(T)}\big) \\ &\leq H\left(\bar{\boldsymbol{G}}_{22}\bar{\boldsymbol{u}}_{22}^{(T)} \;\middle|\; \bar{\boldsymbol{G}}_{10} \begin{pmatrix} \mathbf{0}^{(T)} \\ (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)} \end{pmatrix} \right). \end{aligned}$$

Since $\bar{\boldsymbol{G}}_{mk}$ is lower triangular with nonzero diagonal, it is invertible, implying that

$$\begin{aligned} H\big(\bar{\boldsymbol{s}}_{22}^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big) &\leq H\big(\bar{\boldsymbol{u}}_{22}^{(T)} \bigm| (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\ &= H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{P}})^{(T)} \bigm| (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\ &\leq (n_{22} - n_{12})^+. \end{aligned}$$

Together, this shows that

$$\begin{aligned} T(\bar{R}_{11} + \bar{R}_{12} + \bar{R}_{22} - 2\varepsilon) &\leq T \max\{n_{11}, n_{12}\} + H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{P}})^{(T)} \bigm| (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\ &\leq T(\max\{n_{11}, n_{12}\} + (n_{22} - n_{12})^+). \end{aligned} \tag{77}$$

Therefore, as $T \to \infty$ and $\varepsilon \to 0$, we have (29a). Similarly, we can prove (29b), (29c), and (29d).

We now establish the upper bound (29e). Starting with Fano's inequality,

$$
\begin{aligned}
T(\bar{R}_{11}+\bar{R}_{12}-\varepsilon) &\leq I\big(w_{11}, w_{12}; \bar{\boldsymbol{y}}_1^{(T)}\big) \\
&\leq I\big(w_{11}, w_{12}; \bar{\boldsymbol{y}}_1^{(T)}, \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}\big) \\
&= I\big(w_{11}, w_{12}; \bar{\boldsymbol{y}}_1^{(T)}, \bar{\boldsymbol{s}}_{21}^{(T)} \bigm| w_{21}\big) \\
&= I\big(w_{11}, w_{12}; \bar{\boldsymbol{s}}_{21}^{(T)} \bigm| w_{21}\big) + I\big(w_{11}, w_{12}; \bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}\big) \\
&\leq H\big(\bar{\boldsymbol{s}}_{21}^{(T)} \bigm| w_{21}\big) + H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}\big) - H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}, w_{11}, w_{12}\big) \\
&= H\big(\bar{\boldsymbol{s}}_{21}^{(T)} \bigm| w_{21}\big) + H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}\big) - H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}, w_{11}, w_{12}, \bar{\boldsymbol{u}}_{11}^{(T)}, \bar{\boldsymbol{u}}_{21}^{(T)}\big) \\
&= H\big(\bar{\boldsymbol{s}}_{21}^{(T)} \bigm| w_{21}\big) + H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}\big) - H\big(\bar{\boldsymbol{s}}_{12}^{(T)} \bigm| w_{12}\big). \qquad (78)
\end{aligned}
$$

Similarly, we have

$$
T(\bar{R}_{21}+\bar{R}_{22}-\varepsilon) \leq H\big(\bar{\boldsymbol{s}}_{12}^{(T)} \bigm| w_{12}\big) + H\big(\bar{\boldsymbol{y}}_2^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big) - H\big(\bar{\boldsymbol{s}}_{21}^{(T)} \bigm| w_{21}\big). \qquad (79)
$$

Adding (78) and (79), we obtain

$$
\begin{aligned}
T(\bar{R}_{11}+\bar{R}_{12}+\bar{R}_{21}+\bar{R}_{22}-2\varepsilon) &\leq H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, w_{21}\big) + H\big(\bar{\boldsymbol{y}}_2^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, w_{12}\big) \\
&\leq H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, \bar{\boldsymbol{u}}_{21}^{(T)}\big) + H\big(\bar{\boldsymbol{y}}_2^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, \bar{\boldsymbol{u}}_{12}^{(T)}\big),
\end{aligned}
$$

where in the last line we have used that $\bar{\boldsymbol{u}}_{mk}^{(T)}$ is only a function of w_{mk}. For the first term, we obtain using invertibility of the matrices $\bar{\boldsymbol{G}}_{mk}$,

$$
\begin{aligned}
&H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| \bar{\boldsymbol{s}}_{21}^{(T)}, \bar{\boldsymbol{u}}_{21}^{(T)}\big) \\
&\quad = H\big(\bar{\boldsymbol{y}}_1^{(T)} \bigm| (\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}, \bar{\boldsymbol{u}}_{21}^{(T)}\big) \\
&\quad = H\left(\bar{\boldsymbol{G}}_{11}\begin{pmatrix}(\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}\\(\bar{\boldsymbol{u}}_{11}^{\mathsf{P}})^{(T)}\end{pmatrix} \oplus \bar{\boldsymbol{G}}_{12}\begin{pmatrix}\mathbf{0}^{(T)}\\(\bar{\boldsymbol{u}}_{12}^{\mathsf{C}})^{(T)}\end{pmatrix} \oplus \bar{\boldsymbol{G}}_{10}\begin{pmatrix}\mathbf{0}^{(T)}\\(\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\end{pmatrix} \middle| (\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}\right) \\
&\quad \leq H\left(\bar{\boldsymbol{G}}_{11}\begin{pmatrix}\mathbf{0}^{(T)}\\(\bar{\boldsymbol{u}}_{11}^{\mathsf{P}})^{(T)}\end{pmatrix} \oplus \bar{\boldsymbol{G}}_{12}\begin{pmatrix}\mathbf{0}^{(T)}\\(\bar{\boldsymbol{u}}_{12}^{\mathsf{C}})^{(T)}\end{pmatrix} \oplus \bar{\boldsymbol{G}}_{10}\begin{pmatrix}\mathbf{0}^{(T)}\\(\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\end{pmatrix}\right).
\end{aligned}
$$

Since the matrices $\bar{\boldsymbol{G}}_{mk}$ are lower triangular, this last term is upper bounded by

$$
T\max\{n_{12}, (n_{11}-n_{21})^+\}.
$$

By an analogous argument,

$$
H\big(\bar{\boldsymbol{y}}_2^{(T)} \bigm| \bar{\boldsymbol{s}}_{12}^{(T)}, \bar{\boldsymbol{u}}_{12}^{(T)}\big) \leq T\max\{n_{21}, (n_{22}-n_{12})^+\}.
$$

Together, this shows that

$$
T(\bar{R}_{11}+\bar{R}_{12}+\bar{R}_{21}+\bar{R}_{22}-2\varepsilon) \leq T\big(\max\{n_{12}, (n_{11}-n_{21})^+\} + \max\{n_{21}, (n_{22}-n_{12})^+\}\big),
$$

proving (29e) as $T\to\infty$ and $\varepsilon\to 0$. Similarly, we can prove (29f).

We now establish the bound (29g). By Fano's inequality,

$$
\begin{aligned}
T(\bar{R}_{21}+\bar{R}_{22}-\varepsilon) &\leq I\big(w_{21}, w_{22}; \bar{\boldsymbol{y}}_2^{(T)}\big) \\
&\leq I\big(w_{21}, w_{22}; \bar{\boldsymbol{y}}_2^{(T)}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}, w_{12}\big) \\
&= I\big(w_{21}, w_{22}; \bar{\boldsymbol{y}}_2^{(T)}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)} \bigm| w_{12}\big) \\
&= I\big(w_{21}, w_{22}; (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)} \bigm| w_{12}\big) + I\big(w_{21}, w_{22}; \bar{\boldsymbol{y}}_2^{(T)} \bigm| w_{12}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\
&= H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) + I\big(w_{21}, w_{22}, \bar{\boldsymbol{s}}_{21}^{(T)}; \bar{\boldsymbol{y}}_2^{(T)} \bigm| w_{12}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\
&\quad - I\big(\bar{\boldsymbol{s}}_{21}^{(T)}; \bar{\boldsymbol{y}}_2^{(T)} \bigm| w_{12}, w_{21}, w_{22}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\
&= H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) + H\big(\bar{\boldsymbol{y}}_2^{(T)} \bigm| w_{12}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) - H\big((\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}\big). \qquad (80)
\end{aligned}
$$

Moreover, using again Fano's inequality,

$$\begin{aligned}
T(\bar{R}_{11}-\varepsilon) &\leq I\big(w_{11};\bar{\boldsymbol{y}}_1^{(T)}\big) \\
&\leq I\big(w_{11};\bar{\boldsymbol{y}}_1^{(T)}, w_{12}, w_{21}, w_{22}\big) \\
&= I\big(w_{11};\bar{\boldsymbol{y}}_1^{(T)} \mid w_{12}, w_{21}, w_{22}\big) \\
&= H\big(\bar{\boldsymbol{y}}_1^{(T)} \mid w_{12}, w_{21}, w_{22}\big) \\
&= H\big(\bar{\boldsymbol{u}}_{11}^{(T)}\big) \\
&= H\big((\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}\big) + H\big((\bar{\boldsymbol{u}}_{11}^{\mathsf{P}})^{(T)} \mid (\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}\big).
\end{aligned} \tag{81}$$

Adding (80) and (81) yields

$$\begin{aligned}
& T(\bar{R}_{11}+\bar{R}_{21}+\bar{R}_{22}-2\varepsilon) \\
&\quad \leq H\big(\bar{\boldsymbol{y}}_2^{(T)} \mid w_{12}, (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) + H\big((\bar{\boldsymbol{u}}_{11}^{\mathsf{P}})^{(T)} \mid (\bar{\boldsymbol{u}}_{11}^{\mathsf{C}})^{(T)}\big) + H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\
&\quad \leq T\big(\max\{n_{21},(n_{22}-n_{12})^+\} + (n_{11}-n_{21})^+\big) + H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big).
\end{aligned} \tag{82}$$

Combined with (77) derived earlier, we obtain

$$\begin{aligned}
& T(2\bar{R}_{11}+\bar{R}_{12}+\bar{R}_{21}+2\bar{R}_{22}-4\varepsilon) \\
&\quad \leq T\big(\max\{n_{11},n_{12}\} + \max\{n_{21},(n_{22}-n_{12})^+\} + (n_{11}-n_{21})^+\big) \\
&\quad\quad + H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) + H\big((\bar{\boldsymbol{u}}_{22}^{\mathsf{P}})^{(T)} \mid (\bar{\boldsymbol{u}}_{22}^{\mathsf{C}})^{(T)}\big) \\
&\quad = T\big(\max\{n_{11},n_{12}\} + \max\{n_{21},(n_{22}-n_{12})^+\} + (n_{11}-n_{21})^+\big) + H\big((\bar{\boldsymbol{u}}_{22})^{(T)}\big).
\end{aligned}$$

Since $(\bar{\boldsymbol{u}}_{22})^{(T)}$ is a deterministic function of w_{22}, we have

$$H\big((\bar{\boldsymbol{u}}_{22})^{(T)}\big) \leq H(w_{22}) = T\bar{R}_{22}. \tag{83}$$

From (83), we obtain

$$T(2\bar{R}_{11}+\bar{R}_{12}+\bar{R}_{21}+\bar{R}_{22}-4\varepsilon) \leq T\big(\max\{n_{11},n_{12}\} + \max\{n_{21},(n_{22}-n_{12})^+\} + (n_{11}-n_{21})^+\big).$$

Letting $T\to\infty$ and $\varepsilon\to 0$ yields the upper bound (29g). Similarly, we can prove (29h)–(29j). ■

Remark: Equation 83 is a key step in the derivation of the outer bound (29g). If we had used the standard bound $H\big((\bar{\boldsymbol{u}}_{22})^{(T)}\big) \leq T\max\{n_{22},n_{21}\}$, we would have obtained a looser bound than (29g).

APPENDIX C
ANALYSIS OF MISMATCHED ENCODERS AND DECODERS

The proof of Theorem 6 in Section VI-A assumes that the precise channel gains h_{mk} are available at all encoders and decoders. Here we assume instead that these channel gains are only known approximately at any node in the network. As we will see, the only effect of this change in available channel state information is to decrease the minimum constellation distance seen at the receivers.

Formally, assume both transmitters and receivers have only access to estimates $\hat{h}_{mk}$ of h_{mk} satisfying

$$|h_{mk}-\hat{h}_{mk}| \leq \varepsilon \triangleq 2^{-\max_{m,k} n_{mk}}. \tag{84}$$

In other words, all transmitters and receivers have access to a $\max_{m,k} n_{mk}$-bit quantization of the channel gains. Since we know a-priori that $h_{mk}\in(1,2]$, we can assume without loss of generality that $\hat{h}_{mk}\in(1,2]$ as well.

Each transmitter k forms the modulated symbol u_{mk} from the message w_{mk}. From these modulated signals, the channel inputs

$$\begin{aligned}
x_1 &\triangleq \hat{h}_{22}u_{11} + \hat{h}_{12}u_{21}, \\
x_2 &\triangleq \hat{h}_{21}u_{12} + \hat{h}_{11}u_{22}
\end{aligned}$$

are formed. In other words, the transmitters treat the estimated channel gains $\hat{h}_{mk}$ as if they were the correct ones; the encoders are thus mismatched. The modulation process from w_{1k} to u_{1k} is the same as in the matched case analyzed in Section VI-A. Since $\hat{h}_{mk}^2 \leq 4$ and $u_{mk}^2 \leq 1/16$, the resulting channel input x_k satisfies the unit average power constraint at the transmitters.

The channel output at receiver one is

$$\begin{aligned} y_1 &= 2^{n_{11}} h_{11} x_1 + 2^{n_{12}} h_{12} x_2 + z_1 \\ &= h_{11}\hat{h}_{22} 2^{n_{11}} u_{11} + h_{12}\hat{h}_{21} 2^{n_{12}} u_{12} + \hat{h}_{12} h_{11} 2^{n_{11}} u_{21} + h_{12}\hat{h}_{11} 2^{n_{12}} u_{22} + z_1 \\ &= \big(h_{11}\hat{h}_{22} 2^{n_{11}} u_{11} + h_{12}\hat{h}_{21} 2^{n_{12}} u_{12}^{\mathsf{C}}\big) + \big(\hat{h}_{12} h_{11} 2^{n_{11}} u_{21}^{\mathsf{C}} + h_{12}\hat{h}_{11} 2^{n_{12}} u_{22}^{\mathsf{C}}\big) + \big(h_{12}\hat{h}_{11} 2^{n_{12}} u_{22}^{\mathsf{P}} + z_1\big), \end{aligned}$$

As in the matched case, the received signal consists of desired signals, interference signals, and signals treated as noise. For the third part treated as noise, we have

$$|h_{12}\hat{h}_{11} 2^{n_{12}} u_{22}^{\mathsf{P}}| \leq 1, \tag{85}$$

since

$$2^{n_{12}} u_{22}^{\mathsf{P}} \in [0, 1/4).$$

The demodulator at receiver one searches for

$$\begin{aligned} \hat{s}_{11} &\triangleq 2^{n_{11}} \hat{u}_{11}, \\ \hat{s}_{12} &\triangleq 2^{n_{12}} \hat{u}_{12}^{\mathsf{C}}, \\ \hat{s}_{10} &\triangleq 2^{n_{11}} \hat{u}_{21}^{\mathsf{C}} + 2^{n_{12}} \hat{u}_{22}^{\mathsf{C}} \end{aligned}$$

that minimizes

$$|y_1 - \hat{h}_{11}\hat{h}_{22}\hat{s}_{11} - \hat{h}_{12}\hat{h}_{21}\hat{s}_{12} - \hat{h}_{12}\hat{h}_{11}\hat{s}_{10}|.$$

The demodulator then declares the minimizing $2^{-n_{11}}\hat{s}_{11}$ and $2^{-n_{12}}\hat{s}_{12}$ as the demodulated symbols and discards $\hat{s}_{10}$. Note that the entire demodulation process depends solely on the estimated channel gains $\hat{h}_{mk}$ and not on the actual channel gains h_{mk}. Furthermore, the demodulator is the maximum-likelihood detector only if the estimated channel gains coincide with the actual channel gains. Thus, the demodulator is mismatched.

We now analyze the probability of error of this mismatched demodulator. There are two contributions to this probability of error. One is due to noise, the other one due to mismatched detection. Let d' be the minimum distance between any two noiseless estimated received signals (as assumed by the mismatched demodulator using $\hat{h}_{mk}$). Let $\hat{d}$ be the maximum distance between the noiseless received signal $y_1 - z_1 - h_{12}\hat{h}_{11} 2^{n_{12}} u_{22}^{\mathsf{P}}$ and the estimated received signal with the same channel inputs. The probability of error of the demodulator is then upper bounded by

$$\begin{aligned} \mathbb{P}\big(\hat{u}_{1k} \neq u_{1k} \text{ for } k \in \{1,2\}\big) &\leq 2\mathbb{P}(z_1 + |h_{12}\hat{h}_{11} 2^{n_{12}} u_{22}^{\mathsf{P}}| + \hat{d} \geq d'/2) \\ &\leq 2\mathbb{P}(z_1 \geq d'/2 - \hat{d} - 1), \end{aligned} \tag{86}$$

where we have used (85).

We start by upper bounding the mismatch distance $\hat{d}$. We have

$$\begin{aligned} \hat{d} &\triangleq \max_{(u_{mk})} \big|\hat{h}_{22} 2^{n_{11}} u_{11}(h_{11} - \hat{h}_{11}) + \hat{h}_{21} 2^{n_{12}} u_{12}^{\mathsf{C}}(h_{12} - \hat{h}_{12}) \\ &\qquad\qquad + \hat{h}_{12} 2^{n_{11}} u_{21}^{\mathsf{C}}(h_{11} - \hat{h}_{11}) + \hat{h}_{11} 2^{n_{12}} u_{22}^{\mathsf{C}}(h_{12} - \hat{h}_{12})\big| \\ &\leq 4 \cdot 2 \cdot 2^{n_{11}} \cdot \tfrac{1}{4} \cdot \varepsilon \\ &\leq 2, \end{aligned} \tag{87}$$

where we have used (84), that $|u_{mk}| \leq 1/4$, and that $|\hat{h}_{mk}| \leq 2$.

We continue by lower bounding the distance d' between the estimated received signal (i.e., as assumed by the mismatched detector) generated by the correct (s_{11}, s_{12}, s_{10}) and by any other triple $(s'_{11}, s'_{12}, s'_{10})$. By the triangle inequality,

$$\begin{aligned} d' &\triangleq \min_{\substack{(s_{11},s_{12},s_{10}) \\ \neq (s'_{11},s'_{12},s'_{10})}} \left|\hat{h}_{11}\hat{h}_{22}(s_{11} - s'_{11}) + \hat{h}_{12}\hat{h}_{21}(s_{12} - s'_{12}) + \hat{h}_{12}\hat{h}_{11}(s_{10} - s'_{10})\right| \\ &\geq \min_{\substack{(s_{11},s_{12},s_{10}) \\ \neq (s'_{11},s'_{12},s'_{10})}} \left|h_{11}h_{22}(s_{11} - s'_{11}) + h_{12}h_{21}(s_{12} - s'_{12}) + h_{12}h_{11}(s_{10} - s'_{10})\right| - 3 \cdot 5/2 \\ &\geq d - 8, \end{aligned} \tag{88}$$

where d denotes the minimum distance in the matched case as analyzed in Section VI-A. Here we have used that

$$\begin{aligned} |s_{11} - s'_{11}||\hat{h}_{11}\hat{h}_{22} - h_{11}h_{22}| &\leq 2^{n_{11}-1}|(\hat{h}_{11} - h_{11})(\hat{h}_{22} - h_{22}) + h_{22}(\hat{h}_{11} - h_{11}) + h_{11}(\hat{h}_{22} - h_{22})| \\ &\leq 2^{n_{11}-1} \cdot 5 \cdot \varepsilon \\ &\leq 5/2 \end{aligned}$$

by (84), and similarly for the other two terms.

Combining (87) and (88) shows that

$$\frac{d\prime}{2} - \hat{d} - 1 \geq \frac{d}{2} - 7.$$

By (86), this shows that the probability of demodulation error is upper bounded by

$$2\mathbb{P}(z_1 \geq d/2 - \hat{d} - 1) \leq 2\mathbb{P}(z_1 \geq d/2 - 7). \tag{89}$$

More generally, consider the correct decoding region, and label the incorrect decoding regions in increasing order of distance to the correct one, starting from $\ell \geq 1$ (see Fig. 14 in Section VI-A). Since the minimum distance between any two received signal points is at least d, and taking the mismatch between the decoding metrics into account, the probability of the received signal falling into the ℓth such region is upper bounded by

$$\mathbb{P}(z_1 \geq \ell d'/4 - \hat{d} - 1) \leq \mathbb{P}(z_1 \geq \ell(d - 8)/4 - 3). \tag{90}$$

Appendix D
Proof of Lemma 10 in Section VI-B

The inequalities (48a)–(48f) have been already proved in [10, Lemma 5.2, Theorem 5.3]. Here we present the proof for inequalities (48g)–(48j). First, we establish the bound (48g).

Define $s_{mk}[t]$ as the contribution of transmitter k at receiver m, corrupted by receiver noise $z_m[t]$, i.e.,

$$s_{mk}[t] \triangleq 2^{n_{mk}} h_{mk} x_k[t] + z_m[t]. \tag{91}$$

For block length T, we have,

$$\begin{aligned} T(R_{22} - \varepsilon) &\leq I\big(w_{22}; y_2^{(T)}\big) \\ &\leq I\big(w_{22}; y_2^{(T)}, s_{12}^{(T)}, x_1^{(T)}, w_{12}\big) \\ &= I\big(w_{22}; y_2^{(T)}, s_{12}^{(T)} \mid x_1^{(T)}, w_{12}\big) \\ &= I\big(w_{22}; s_{22}^{(T)}, s_{12}^{(T)} \mid w_{12}\big) \\ &= I\big(w_{22}; s_{12}^{(T)} \mid w_{12}\big) + I\big(w_{22}; s_{22}^{(T)} \mid s_{12}^{(T)}, w_{12}\big) \\ &= h\big(s_{12}^{(T)} \mid w_{12}\big) - h\big(z_1^{(T)}\big) + h\big(s_{22}^{(T)} \mid s_{12}^{(T)}, w_{12}\big) - h\big(z_2^{(T)}\big), \end{aligned} \tag{92}$$

where the first step follows from Fano's inequality. Again from Fano's inequality, we have,

$$\begin{aligned} T(R_{11}+R_{12}-\varepsilon) &\le I\big(w_{11},w_{12};y_1^{(T)}\big) \\ &\le I\big(w_{11},w_{12},w_{21};y_1^{(T)}\big) \\ &= h\big(y_1^{(T)}\big)-h\big(y_1^{(T)} \mid w_{11},w_{12},w_{21}\big) \\ &= h\big(y_1^{(T)}\big)-h\big(s_{12}^{(T)} \mid w_{12}\big). \end{aligned} \tag{93}$$

Adding (92) and (93) yields

$$T(R_{11}+R_{12}+R_{22}-2\varepsilon) \le h\big(y_1^{(T)}\big)-h\big(z_1^{(T)}\big)+h\big(s_{22}^{(T)} \mid s_{12}^{(T)},w_{12}\big)-h\big(z_2^{(T)}\big). \tag{94}$$

Using Fano's inequality at receiver two, we have

$$\begin{aligned} T(R_{21}+R_{22}-\varepsilon) &\le I\big(w_{21},w_{22};y_2^{(T)}\big) \\ &\le I\big(w_{21},w_{22};y_2^{(T)},s_{12}^{(T)},w_{12}\big) \\ &= I\big(w_{21},w_{22};y_2^{(T)},s_{12}^{(T)} \mid w_{12}\big) \\ &= I\big(w_{21},w_{22};s_{12}^{(T)} \mid w_{12}\big)+I\big(w_{21},w_{22};y_2^{(T)} \mid s_{12}^{(T)},w_{12}\big) \\ &= h\big(s_{12}^{(T)} \mid w_{12}\big)-h\big(z_1^{(T)}\big)+h\big(y_2^{(T)} \mid s_{12}^{(T)},w_{12}\big)-h\big(s_{21}^{(T)} \mid w_{21}\big). \end{aligned} \tag{95}$$

Moreover, Fano's inequality at receiver one yields

$$\begin{aligned} T(R_{11}-\varepsilon) &\le I\big(w_{11};y_1^{(T)}\big) \\ &\le I\big(w_{11};y_1^{(T)},s_{21}^{(T)},w_{12},w_{21},w_{22}\big) \\ &= I\big(w_{11};y_1^{(T)},s_{21}^{(T)} \mid w_{12},w_{21},w_{22}\big) \\ &= I\big(w_{11};s_{11}^{(T)},s_{21}^{(T)} \mid w_{12},w_{21},w_{22}\big) \\ &= h\big(s_{11}^{(T)},s_{21}^{(T)} \mid w_{12},w_{21},w_{22}\big)-h\big(s_{11}^{(T)},s_{21}^{(T)} \mid w_{11},w_{12},w_{21},w_{22}\big) \\ &= h\big(s_{21}^{(T)} \mid w_{21}\big)+h\big(s_{11}^{(T)} \mid s_{21}^{(T)},w_{21}\big)-h\big(z_1^{(T)},z_2^{(T)}\big). \end{aligned} \tag{96}$$

Adding (95) and (96) yields

$$\begin{aligned} T(R_{11}&+R_{21}+R_{22}-2\varepsilon) \\ &= h\big(s_{12}^{(T)} \mid w_{12}\big)-h\big(z_1^{(T)}\big)+h\big(y_2^{(T)} \mid s_{12}^{(T)},w_{12}\big)+h\big(s_{11}^{(T)} \mid s_{21}^{(T)},w_{21}\big)-h\big(z_1^{(T)},z_2^{(T)}\big). \end{aligned} \tag{97}$$

Adding (97) and (94) derived earlier, we obtain

$$\begin{aligned} T(2R_{11}&+R_{12}+R_{21}+2R_{22}-4\varepsilon) \\ &\le h\big(y_1^{(T)}\big)-2h\big(z_1^{(T)}\big)+h\big(s_{22}^{(T)} \mid s_{12}^{(T)},w_{12}\big)-h\big(z_2^{(T)}\big)+h\big(s_{12}^{(T)} \mid w_{12}\big) \\ &\quad +h\big(y_2^{(T)} \mid s_{12}^{(T)},w_{12}\big)+h\big(s_{11}^{(T)} \mid s_{21}^{(T)},w_{21}\big)-h\big(z_1^{(T)},z_2^{(T)}\big) \\ &= h\big(y_1^{(T)}\big)-2h\big(z_1^{(T)}\big)+h\big(s_{22}^{(T)},s_{12}^{(T)} \mid w_{12}\big)-h\big(z_2^{(T)}\big) \\ &\quad +h\big(y_2^{(T)} \mid s_{12}^{(T)},w_{12}\big)+h\big(s_{11}^{(T)} \mid s_{21}^{(T)},w_{21}\big)-h\big(z_1^{(T)},z_2^{(T)}\big). \end{aligned} \tag{98}$$

Since

$$\begin{aligned} h\big(s_{22}^{(T)},s_{12}^{(T)} \mid w_{12}\big)-h\big(z_1^{(T)},z_2^{(T)}\big) &= I\big(w_{22};s_{22}^{(T)},s_{12}^{(T)} \mid w_{12}\big) \\ &\le H(w_{22}) \\ &= TR_{22}, \end{aligned} \tag{99}$$

we obtain from (98) that

$$
\begin{aligned}
&T(2R_{11}+R_{12}+R_{21}+R_{22}-4\varepsilon)\\
&\qquad\qquad \le h\big(y_1^{(T)}\big)-2h\big(z_1^{(T)}\big)-h\big(z_2^{(T)}\big)+h\big(y_2^{(T)} \mid s_{12}^{(T)}, w_{12}\big)+h\big(s_{11}^{(T)} \mid s_{21}^{(T)}, w_{21}\big)\\
&\qquad\qquad \le h\big(y_1^{(T)}\big)-h\big(z_1^{(T)}\big)+h\big(y_2^{(T)} \mid s_{12}^{(T)}\big)-h\big(z_2^{(T)}\big)+h\big(s_{11}^{(T)} \mid s_{21}^{(T)}\big)-h\big(z_1^{(T)}\big)\\
&\qquad\qquad \le \frac{T}{2}\log\left(1+2^{2n_{11}}h_{11}^2+2^{2n_{12}}h_{12}^2\right)+\frac{T}{2}\log\left(1+2^{2n_{21}}h_{21}^2+\frac{2^{2n_{22}}h_{22}^2}{1+2^{2n_{12}}h_{12}^2}\right)\\
&\qquad\qquad\quad +\frac{T}{2}\log\left(1+\frac{2^{2n_{11}}h_{11}^2}{1+2^{2n_{21}}h_{21}^2}\right),
\end{aligned}
$$

where the last inequality follows from the fact that i.i.d. Gaussian random variables maximize conditional differential entropy (see, e.g., [27, Lemma 1]). Letting $T \to \infty$ and $\varepsilon \to 0$ now proves (48g). Inequalities (48h)–(48j) can be proved similarly. ■

Remark: We point out that, as in the deterministic case, (99) is a key step to the derivation of the outer bound for the Gaussian X-channel.

APPENDIX E
PROOF OF LEMMA 13 IN SECTION VII-B

By Fubini's theorem, we have

$$
\begin{aligned}
\mu_4(B) &= \int_{h_{11}=1}^{2}\int_{h_{12}=1}^{2}\int_{h_{21}=1}^{2}\int_{h_{22}=1}^{2} \mathbb{1}_B(h_{11},h_{12},h_{21},h_{22})dh_{22}dh_{21}dh_{12}dh_{11}\\
&= \int_{h_{11}=1}^{2}\int_{h_{12}=1}^{2}\int_{h_{21}=1}^{2}\int_{h_{22}=1}^{2} \mathbb{1}_{\tilde{B}}(h_{11}h_{12},h_{11}h_{22},h_{12}h_{21})dh_{22}dh_{21}dh_{12}dh_{11}\\
&\le \int_{h_{11}=1}^{2}\int_{g_0=1}^{4}\int_{g_1=1}^{4}\int_{g_2=1}^{4} \mathbb{1}_{\tilde{B}}(g_0,g_1,g_2)g_0^{-1}h_{11}^{-1}dg_2dg_1dg_0dh_{11}\\
&\le \int_{h_{11}=1}^{2} \mu_3(\tilde{B})dh_{11}\\
&\le \delta,
\end{aligned}
$$

proving the lemma. ■

ACKNOWLEDGMENT

The authors would like to thank G. Kramer for pointing them to Slepian's 1974 Shannon Lecture [16].

REFERENCES

[1] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3457–3470, Aug. 2008.

[2] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[3] A. S. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *arXiv:0908.2282 [cs.IT]*, Aug. 2009. Submitted to IEEE Trans. Inf. Theory.

[4] C. Suh and D. Tse, "Interference alignment for cellular networks," in *Proc. Allerton Conf.*, pp. 1037–1044, Sept. 2008.

[5] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Proc. Allerton Conf.*, pp. 1188–1195, Sept. 2010.

[6] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment," in *Proc. IEEE ISIT*, pp. 2756–2760, July 2011.

[7] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of three-user MIMO broadcast channel with delayed CSIT," in *Proc. IEEE ISIT*, pp. 209–213, July 2011.

[8] U. Niesen and P. Whiting, "The degrees of freedom of compute-and-forward," *arXiv:1101.2182 [cs.IT]*, Jan. 2011. Submitted to IEEE Trans. Inf. Theory.

[9] G. Bresler, A. Parekh, and D. N. C. Tse, “The approximate capacity of the many-to-one and one-to-many Gaussian interference channels,” *IEEE Trans. Inf. Theory*, vol. 56, pp. 4566–4592, Sept. 2010.

[10] C. Huang, V. R. Cadambe, and S. A. Jafar, “On the capacity and generalized degrees of freedom of the X channel,” *arXiv:0810.4741 [cs.IT]*, Oct. 2008.

[11] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Trans. Inf. Theory*, Apr. 2011.

[12] C. Suh and D. N. C. Tse, “Feedback capacity of the Gaussian interference channel to within 2 bits,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 2667–2685, May 2011.

[13] I.-H. Wang and D. N. C. Tse, “Interference mitigation through limited transmitter cooperation,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 2941–2965, May 2011.

[14] I.-H. Wang and D. N. C. Tse, “Interference mitigation through limited receiver cooperation,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 2913–2940, May 2011.

[15] M. A. Maddah-Ali and D. Tse, “Interference neutralization in distributed lossy source coding,” in *Proc. IEEE ISIT*, pp. 166–170, June 2010.

[16] D. Slepian, “On bandwidth,” *Proc. IEEE*, vol. 64, pp. 292–300, Mar. 1976.

[17] R. H. Etkin and E. Ordentlich, “The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients,” *IEEE Trans. Inf. Theory*, vol. 55, pp. 4932–4946, Nov. 2009.

[18] A. A. El Gamal and M. H. M. Costa, “The capacity region of a class of deterministic interference channels,” *IEEE Trans. Inf. Theory*, vol. 28, pp. 343–346, Mar. 1982.

[19] G. Bresler and D. Tse, “The two-user Gaussian interference channel: a deterministic view,” *Europ. Trans. Telecommun.*, vol. 19, pp. 333–354, Apr. 2008.

[20] M. Anand and P. R. Kumar, “A digital interface for Gaussian relay and interference networks: Lifting codes from the discrete superposition model,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 2548–2564, May 2011.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[22] R. Durret, *Probability: Theory and Examples*. Duxbury Press, third ed., 2004.

[23] V. G. Sprindžuk, *Metric Theory of Diophantine Approximations*. Wiley, 1979.

[24] M. M. Dodson, “Diophantine approximation, Khintchine’s theorem, torus geometry and Hausdorff dimension,” *Sémin. Congr.*, vol. 19, pp. 1–20, 2009.

[25] O. Ordentlich and U. Erez, “Interference alignment at finite SNR for time-invariant channels,” *arXiv:1104.5456 [cs.IT]*, Apr. 2011.

[26] Y. Wu, S. Shamai, and S. Verdú, “Degrees of freedom of the interference channel: A general formula,” in *Proc. IEEE ISIT*, pp. 1362–1366, July 2011.

[27] J. A. Thomas, “Feedback can at most double Gaussian multiple access channel capacity,” *IEEE Trans. Inf. Theory*, vol. 33, pp. 711–716, Sept. 1987.